INFRASTRUCTURE NETWORK ENHANCEMENT INSPIRED BY NATURE

A Dissertation
Presented to
The Academic Faculty

By

L. Fernando Patino Ramirez

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
College of Engineering
Department of Civil and Environmental Engineering

Georgia Institute of Technology

December 2020

© L. Fernando Patino Ramirez 2020

INFRASTRUCTURE NETWORK ENHANCEMENT INSPIRED BY NATURE

Thesis committee:

Dr. Chloé Arson
Department of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Duen Horng Chau
School of Computational Science & Engineering
Georgia Institute of Technology

Dr. David Frost
Department of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Bernardo Caicedo
Department of Civil and Environmental Engineering
Universidad de los Andes

Dr. Sheng Dai
Department of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Gioacchino Viggiani
School of Civil Engineering and Geomechanics
Universite Grenoble Alpes

Date approved: November 6, 2020

If I have seen further, it is by standing on the shoulders of giants.

Sir. Isaac Newton

For Nancy and Fernando

ACKNOWLEDGMENTS

This thesis would not be possible without the help and support of many colleagues, family and friends, to them I owe not only my work, but also my growth as a researcher and a person. Thank you all and muchas gracias.

Starting with my thesis committee, I thank each member for their guidance and knowledge. My advisor, Prof. Chloé Arson, thank you so much for guiding me through the process, for believing in me and my crazy ideas and for making me feel supported at every step of the process. to Prof. Bernardo Caicedo, for encouraging me to get a PhD in the first place and for being a permanent source of wisdom. Prof. David Frost, for always reminding me to focus on the bigger picture, teaching me about topics beyond geotech, and for adopting me as part of his group at many times. To Prof. Cino Viggiani, from whom I learned to be precise and to ask the right questions, for hosting me in France, and for all the pleasant conversations. Prof. Sheng Dai, for his support to my thesis, everything I learned from him, and the support to Geo-Society. And last but not least Prof. Polo Chau, for helping me dive into computer science and for his contributions to my thesis work.

To my friends in Atlanta, Geo's and not-geo's. I have a deep appreciation for my Geo-friends and colleagues, I thank them for being an inspiring group to work with, but also for being supportive and present along the way, not only to have technical discussions, but more importantly to have fun and also to vent when needed. The list includes but is not limited to: Rod, Mahdi, Nimishita, Karie, Xianda, Koochul, Tingting, Albert, Wilson, Boyoung, Andrés, Seth, Frikret and Xenia. I also thank my friends outside Geosystems and outside GT, they kept me grounded thorough the process, significantly contributed to my work-life balance during the PhD, and taught me about other cultures, food, dance styles, languages, and academic and non-academic fields; to all of you I say thanks. Lina, for being extremely supportive, caring and patient, to Oliver for being my confident, colleague and friend, Leo for being always there with a different and valuable point of view, Juli

for being a great roommie, friend and admirable professional, Carlos for his patience and sense of humor, Asta for teaching me about football and being one of the kindest people I have met, Diego for his good heart, and many others that, if listed, would put in doubt the amount of time I actually spent working on my thesis.

Last but not least, I want to thank my family. Mis papás, a quien debo todo lo que soy y todo lo que seré, a mi papá que no pudo leer esto, pero siempre estuvo presente en el proceso y estará presente en mi vida, mi mamá por ser el apoyo constante y una verdadera motivación para siempre lograr más, y a Cami por siempre estar ahí. To my family in the US, for their immense support, for bearing with me all these years, and for making me feel part of the family as if we had never been apart on opposite coasts, thanks Cami, Manuel, Sebas, Martha, Ricardo and AJ. Muchas gracias.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

The exploration of the underground space is essential to the exploitation of energy resources, the foundation and design of structures and, more importantly, to the sustained development of cities both in terms of transportation and access to utilities. The urbanization and densification of the population around the globe has caused a reduction of the above-ground space, and an increased demand for urban transportation and utility systems. Such increased demand and low availability of above-ground space has highlighted the social and environmental benefits that underground construction offers, and the need for new and more efficient subsurface design methods.

Nature has created and perfected exceptional, sustainable and cost-efficient strategies to build underground cavities (e.g., worms and ants), and transportation networks (e.g., roots, leaves and fungal networks). Underground networks are global, interconnected systems, subjected to complex environmental constrains that vary in space and time. Taking inspiration from natural systems, we analyze the mechanical performance of underground structures at the scale of the cavity, and we optimize cost and resiliency of the scale of the entire network.

In the first part of the thesis, we take inspiration from biology to optimize networks that connect resources spread over a finite domain and we apply biological network models to the design of engineered systems. We first study the morphogenesis of slime mold, and we describe its behavior with transport indirect graphs. We apply slime mold underlying network deployment strategies to the optimization of fracture patterns in rock. Then, we use a leaf venation inspired algorithm to model roads and public transportation, benchmarking its performance against theoretical bounds and applying it to the design of transportation networks in the city of Atlanta.

In the second part of the thesis, we focus on the mechanical behavior of underground cavities embedded in granular soil at shallow depth, which is relevant to in-situ testing,

tunneling, anchoring and arching. We focus on the mechanical behavior and failure patterns of such cavities, and we examine common assumptions used in practice to model cavity expansion. We first present the results of a comprehensive experimental study on pressurized vessels under geo-static stress. We analyze the influence of the length of the cavity, the density of the soil and the vertical surcharge load on failure mechanisms and strain patterns. We then use the finite element method (FEM) and machine learning (ML) algorithms to quantify the influence of various cavity and soil parameters and understand their influence on the behavior of the cavity. Lastly, we show an application to real design problems, using an ant colony optimization algorithm to automate and accelerate the design process of horizontal directional drills.

CHAPTER 1
INTRODUCTION

The exploration of the underground space is essential to the exploitation of energy resources, foundation and design of structures and infrastructure and, more importantly, to the sustained development of cities both in terms of transportation and access to utilities. The urbanization and densification of the population around the globe has caused a reduction of the above-ground space, in addition to an increased demand for urban transportation and utility systems, including but not limited to: train, water, sewage, power, data and gas. Such increased demand and low availability of above-ground space highlights the key social and environmental benefits that underground construction offers, and the need for new subsurface design methods [1].

As urban settlements and population become larger, systems must accommodate not only to an increasing and varying demand (and thus capacity), but also to a suitable distribution of resources. Transport of such resources have to guarantee an adequate level of service, and are thus becoming more interconnected and complex. Civil infrastructure is traditionally designed to optimize objective functions such as traffic flow and power grid service under static constraints, such as land use regulations and geometrical compatibility with existing infrastructure. Nonetheless, the dynamic nature of urban networks has brought to light the need for design methods that can adapt to volatile and uncertain constraints and demand over space and time. Moreover, the interconnected nature of infrastructure networks with the environment, and global threats such as pandemics, climate change and terrorism, also call for resilient systems that can sustain damage and adapt to disruptions, maintaining an adequate level of service [2].

In order to overcome the current limitations in the field, a change of paradigm has to take place in the way underground networks, and transportation systems in general,

are conceived. Two significant obstacles are that: (i) underground construction is usually seen as a last resource, once above-ground options have been exhausted [3], disregarding indirect, non-financial benefits; and (ii) projects are conceived incrementally, despite the fact that they belong to an interconnected network where local changes result in global effects. If underground networks are conceived as global systems, subject to complexities that span over space and time, new, more comprehensive methods will be possible.

Recent awareness in the value of the externalities associated with urban underground space (UUS) has improved the understanding of the trade-offs, long term benefits and threats of underground networks. For instance, a methodology was proposed [4] to quantify the value of underground space, to better guide design decisions and inform public policy; similarly, an approach [5] to value the UUS using Geographic Information System (GIS) and mathematical concepts was developed to better drive regulatory planning and master plans in urban regions.

A common way to spark new ideas to pursue the highly needed new methods and tools, is the creation of similes between the sought-after function and comparable strategies validated in other fields or with other purposes. Since the beginning of humanity, as the species began settling, our ancestors initially copied strategies used by nature to create tools, protect themselves and acquire resources [6]. Then, humans started to mimic strategies from nature instead of copying them, creating new tools that were used to solve the same problems with different methods [7]. For instance, humans initially used untreated fur and pelts from animals, then started to treat them and eventually created substitutes like textiles, that mimic the functionality of fur, using different, more accessible methods.

This process of using phenomena from nature to stimulate the development of non-biological science and technology is known as bio-inspiration or biomimicry [8], and will be the focus of the present thesis. A related field, named bio-mediation, aims to apply bio-chemical processes to achieve a desired non-biological function, for example, using microbially mediated calcite precipitation to heal cracks in cementicious materials [9].

Bio-inspired design usually follows three steps. First, the observation of a strategy or function displayed by an organism in nature inspires the application, by abstraction to a 'human' need. Then, the fundamental processes behind the biological strategy are modeled. Lastly, the given strategy is mapped and transferred to the 'human' application, by using the technology that is adequate or available [10]. Therefore, the implemented bio-inspired solution does not necessarily replicate or precisely clones the same fundamental mechanisms of the biological system (nor does it intend to).Bio-inspired design sometimes follows a reverse process, in which designers turn to nature to identify problems that are simimar to the ones that they are facing in engineering. Once a suitable biological candidate is found, the solution is adapted, as explained above. There is no shortage of examples of bio-inspired technologies, all the way from claws and knives, to the movement of the moon and stars that inspired the formulation of the gravitation law, to the flight of birds that inspired the creation of airplanes [6, 11].

Even though these principles of bio-inspiration have been present through the history of humankind and the development of science [7], the concept was not formalized until the early 1950s, when Otto Schmitt coined the term as 'the biological approach to engineering' [12]. Then, the term 'bionics' was introduced by Jack Steel of the US Air force, defining it as 'copying and taking ideas from nature', and it became popular thanks to the book: Bionics: The Science of Living Machines [13]. Nonetheless, it was not until the early 1990's that a significant amount of scientific literature addressing bio-inspiration started to surface, sparking at the same time the curiosity of academia, and the need to create dedicated journals and conferences for these studies that are by definition interdisciplinary, and therefore somehow difficult to fit into journals or professional societies [14]. More recently, in the early 2000's, an increased number of dedicated issues in journals [8], journals entirely dedicated to bio-inspiration [15], and in general an increased awareness and acceptance of the field [14], resulted in a significant amount of studies [12].

The field of geotechnics, which deals with geo-materials, their mechanical improve-

ment, and the structures that interact with them (foundations, tunnels,...), has explored bio-inspired and bio-mediated mechanisms over the last 15 years, with promising results and methods to improve the state-of-the-art [16]. Such developments are divided into: a) bio-inspiration (or biomimicry): taking inspiration from nature to develop geotechnical engineering solutions that do not necessarily follow the same physical principles, for instance, creating metal or concrete piles that loosely follow the geometry of root systems to maximize anchoring capacity [17]; and b) bio-mediation: applying bio-chemical processes to achieve a desired geotechnical function, for example, using microbially mediated calcite precipitation to improve the shear strength of soils [10]. Among these two, bio-mediated geotechnics has been the main branch of research over the last decade, with widespread industry implementation, the most popular use being MICP (microbially induced calcite precipitation) [18].

On the other hand, bio-inspired geotechnics, in which this thesis falls in, has focused mainly in three main fields of study: foundations, frictional interfaces and burrowing/excavation. Bio-inspired foundations typically assimilate the anchoring mechanisms of root systems in order to apply them to engineered foundation systems, maximizing their anchoring performance [19, 17]. Studies on frictional interfaces have focused on the effect of anisotropy, which is crucial for the soil-structure interaction of buried elements such as retaining walls, piles and footings, and has been inspired by the principles of locomotion of sharks and snakes, favoring a specific direction of movement and increasing resistance in the opposite direction [20, 21]. Lastly, subsurface exploration has motivated the study of excavation mechanisms, for instance, taking inspiration from the efficient excavation mechanisms followed by ants [22], and studying the burrowing processes of animals, such as worms and razor clams, in order to develop self-burrowing probes for soil exploration and characterization [23, 24].

Transportation networks are no exception to the applicability of bio-inspiration and bio-mimicry. Deploying an efficient and mechanically stable network in a constrained

environment, with a finite amount of resources, is a common objective to both biological and human-made systems [25]. Thus, every organism in nature, even the most seemingly simple ones, have developed intricate strategies to adapt to their environment, sometimes in a strikingly similar way to human algorithms and design techniques. This fact has not only improved our understanding of the strategies of nature, but more importantly, has inspired humanity to develop new methods and tools, inspired by nature, exploiting the underlying similarities in the objective functions and constraints to which both biological and human-made transport systems are subjected [25, 26].

Regardless of scale, ranging from from ecosystems and communities such as ant colonies, to simple organisms and organs such as neural arbors and vascular systems, it has been observed that network deployment strategies consist in maximizing travel efficiency and/or redundancy (and therefore resiliency) while minimizing cost. Different trade-offs are achieved between these characteristics depending on the type of network and its functionality. This fact is oftentimes forgotten in engineering practice; and even though the concept of complex systems has been addressed widely, it becomes clear that many of our structures lack the ability to perform well beyond their primary function, in part because many engineering structures are designed as stand-alone elements that are in reality part of complex interconnected networks.

The present thesis focuses on the bio-inspired design of engineering infrastructure networks, with emphasis on underground cavities, and is presented in two parts: (i) Bio-inspired optimization of transport networks, treating the optimal topology of the system as a whole, and (ii) The mechanics of underground cavities in granular soil, focusing on the mechanical behavior and stability of such underground cavities. These two complementary parts provide functional requirements necessary to ensure that the underground networks provide an adequate level of service and performance.

Figure 1.1 shows a graphical representation of the different topics presented in this thesis, where each one of the parts starts with developments in the fundamental knowledge

plane of infrastructure networks (part 1) and underground cavities (part 2). Then, following chapters on each part continue exploring enabling technologies and system integration techniques that test applications and implications of the fundamental knowledge, and at the same time span from the global (network) scale of analysis, to the intermediate (cavity alignment) and local (cross-section) scales of underground networks.

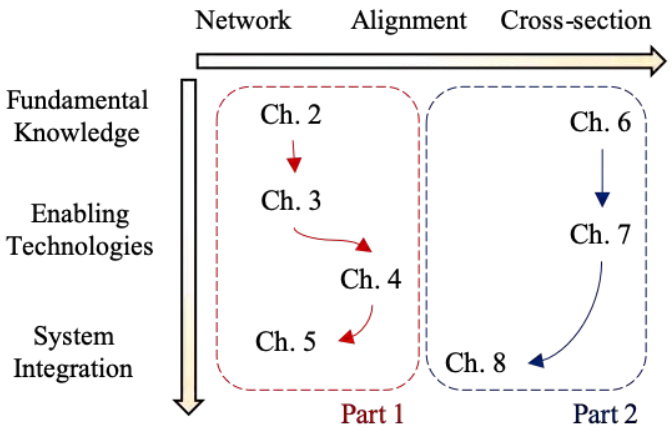

Figure 1.1: Thesis outline. Scheme showing the different scales (network to cross-section) and planes of research (fundamental knowledge to system integration) that the present thesis spans.

In the first part, we take inspiration from natural systems which exhibit outstanding abilities to connect and reach resources spread through a given domain. Biological networks either forage and explore an unknown domain or optimize flow and connectivity within a constrained domain of relatively known characteristics. Among the foraging organisms, we find plant roots, mycelial networks, ant colonies and slime molds. Some constrained networks include leaf venations, vascular systems and neural arbors. We study both types of network in this thesis. First, we present our work on slime mold, its morphogenesis in chapter 2, published in [27], and its behavior as transport indirect graphs in chapter 3, published in [28]). We then apply underlying network deployment strategies to the optimization of fracture patterns in rock in chapter 4, as previously published in [29]. Then, we focus on the application of leaf venation (LV) algorithm to human-made transportation net-

works, namely, roads and public transportation, with a study initially published in [26] in chapter 5. In this study, we benchmark the performance of LV networks against theoretical bounds and apply sit to the design of transportation networks in the city of Atlanta.

The second part of this thesis focuses on the mechanical behavior of underground cavities embedded in granular soil at shallow depth. This problem and specific loading conditions are relevant to a large number of applications, including but not limited to: in-situ testing [30, 31] including pressuremeter testing (PMT) [32], the cone penetration test (CPT) [33, 34] and the dilatometer test (DMT) [35]; tunneling, including micro-tunneling and horizontal directional drilling (HDD) [36, 37, 38]; and anchoring and arching [39, 40, 41]. The goal of this section is to deepen the understanding of the mechanical behavior and failure patterns of such cavities, validate common assumptions used in practice and propose new methods for optimization and design of such cavities. First, we present the results of a comprehensive experimental study on pressurized vessels under biaxial stress published in [42] and presented here in chapter 6. We focus on the failure mechanism and strain patterns and we study the influence of the length of the cavity, the density of the soil and vertical surcharge load on the response of the soil mass. chapter 7 extends such findings using the finite element method (FEM) and machine learning (ML) algorithms to quantify the influence of various cavity and soil parameters, and understand their influence on the behavior of the cavity. The first part of this chapter focuses on the pressure-controlled expansion of the cavities, published on [43], while the second part shows the initial findings on the study of displacement-controlled expansion. Lastly, and application to real design problems is shown in chapter 8, where an ant colony optimization (ACO) algorithm is used to automate and accelerate the design of the geometric alignment of shallow tunnels, specifically, horizontal directional drills, this study has been published in [44].

CHAPTER 2

INFLUENCE OF SUBSTRATE ON SLIME MOLD MORPHOGENESIS

2.1 Aim and summary

In this chapter, we study the behavior of slime mold (*Physarum polycephalum*), a unicellular organism that develops by forming networks. We use slime mold as a model organism to study network resiliency and adaptation in nature. Specifically, this chapter focuses on the influence of the nutritive vs. adverse nature of the substrate on slime mold growth patterns, called morphogenesis. The present study has been published in [27].

Slime mold is a giant single-celled organism that can grow to cover several square meters, forming search fronts that are connected to a system of intersecting veins. An original experimental protocol allowed tracking the shape of slime mold placed in homogeneous substrates containing an attractant nutrient (glucose) or a repellent (salt), or in homogeneous substrates that contained an attractive spot (glucose), an eccentric slime mold and a repulsive spot (salt) in between.

For the first time, the rate of exploration of unexplored areas (primary growth) and the rate of extension in previously explored areas (secondary growth) were rigorously measured, by means of a sophisticated image analysis program. Results show that the chemical composition of the substrate has more influence on the morphology and growth dynamics of slime mold than that of concentrated spots of chemicals. It was also found that on a repulsive substrate, slime mold exhibits a bias towards secondary growth, which suggests that the mucus produced during slime mold migration acts as a protective shell in adverse environments.

2.2 Introduction

Large-scale spatial patterns in biology are common and knowing how these patterns evolve and what their functional role is, can shed light on the evolution of biocomplexity (see e.g. [45, 46, 47, 48]. Morphogenesis has been studied in length at the cell level (see e.g [49, 50, 51, 52]; cells are highly sensitive to geometrical and mechanical constraints from their microenvironment and respond to these conditions by changing shape (see e.g. [53, 54]); these transformations impact cell migration and growth (see e.g. [50, 55, 56, 57]. Cellular migration is a fundamental property of every cell and it is crucial for the development and morphogenesis of animal body plans and organ systems (see e.g. [58, 59, 60]. Cell migration is either in a random direction or directed towards localized cues [61, 62, 63, 64]. Mechanisms of cellular movement have been mostly studied in chemotactic cells, such as neutrophils [61], bacteria [65], Ciliata [66], fungi [67] and cellular slime molds [63].

Due to its extremely fast migration rate and highly irregular shape, the acellular slime mold *Physarum Polycephalum* represents a prime example of differentiated growth and thus offers an attractive model for the analysis of morphogenesis dynamics underlying cellular migration and exploration [68, 69, 70, 71, 72]. *P polycephalum* is a giant single-celled organism that can grow to cover several square meters. Its morphology includes search fronts that are connected to a system of intersecting veins, in which oscillatory flows of the protoplasm, called "shuttle streaming," take place. This vein network allows 1) an efficient distribution of chemical signals, oxygen, nutrients over large distances and 2) cell migration at a speed of few centimeters per hour [73, 74]. The driving force for this protoplasm streaming is a periodic, peristaltic contraction and relaxation of the veins due to the actin-myosin interaction, which is regulated by oscillations of intracellular chemicals such as calcium [75, 76, 77]. As it explores its environment, the slime mold extends temporary arm-like projections named pseudopods. It also secretes continuously a thick extracellular slime [78]. The glycoprotein nature of the extracellular slime coat endows *P polycephalum*

with unique protective and structural properties that favor survival of the migrating, naked slime mold [79]. As the slime mold is foraging, it avoids areas covered with this mucus, which marks previously explored areas [80, 81].

In the presence of chemical substances in the environment, *P. Polycephalum* shows directional movements towards or away from the stimuli (i.e. chemotaxis). *Physarum* morphology, evolution and behaviors are strongly affected by the availability, location and concentration of nutrients. When the slime mold senses attractants (e.g. food cues) using specific receptors located on the membrane, the oscillation frequency in the pseudopod closest to the attractant increases, causing cytoplasm to flow towards the attractant [82]. On the contrary, when repellents such as salts are sensed, the oscillation frequency decreases and the slime mold moves away from the repellent [82]. Although slime molds lack the complex hardware of animals with brains, they live in environments that are as complex and they face the same decision-making challenges [83]. Hence, acellular slime molds have been the subject of a wide range of studies showing that they can solve complex biological and computational problems without any specialized nervous tissue [84, 85, 86, 87, 88, 80, 89, 81, 90, 25].

In this chapter, the objectives are to characterize the morphology and dynamics of *Physarum* exploring various environments. First, we investigate how movement is affected by homogeneous environmental conditions: adverse environment (using salt as a repellent [91]; and nutritive environments (using glucose as a chemo-attractant [92, 76] with 2 different concentrations) and a neutral environment (using plain agar). Second, we analyze the geometrical evolution of slime molds placed at a distance from a nutritive spot (glucose), with and without a repelling spot (salt) in between. We characterize slime molds' movement both temporally and spatially, to capture the full dynamics. To this aim, we develop a program that automatically analyzes sequences of images to track the areas covered and explored by the slime mold, the slime mold shape, the refinement and secondary growth cycles, as well as the distance to the nutritive spot.

2.3 Methods

2.3.1 Species

Physarum Polycephalum, also known as the true slime mold, belongs to the Amoebozoa, the sister group to fungi and animals [93]. Slime molds are found on organic substrates where they feed on microorganisms such as bacteria or fungi [93]. The vegetative form of *P. polycephalum*, the plasmodium, is a vast multinucleate cell that can grow to cover up to a few square meters and crawl at speeds from 0.1 to few centimeters per hour [73, 74]. When hygrometry and food availability decrease, the plasmodium turns into an encysted resting stage made of desiccated spherules called sclerotium [73].

2.3.2 Rearing conditions

Experiments were initiated with a total of 10 sclerotia per strain (Southern Biological, Victoria, Australia). We cultivated slime molds on a 1% agar medium with rolled oat flakes, slime molds were fed every day and the medium was replaced daily. Slime molds were 2 weeks old when the experiment started. All experiments were carried out in the dark at 25 ∘ C temperature and 70% humidity, and ran for 35 h. Pictures were taken with a Canon 70D digital camera.

2.3.3 Experimental setup

Initially we monitored the exploration movement evoked in slime molds in a homogeneous environment. Each slime mold was placed in the center of a circular arena (14.5cm in diameter) with a layer of agar (1% in water) mixed with non-nutritive cellulose (5%). Adding cellulose to the agar mix proved to be useful to obtain a homogeneous pigmentation and to enhance the color contrast between the substrate and slime mold, therefore improving the identification process. A circular hole (2.5cm in diameter) was punched and replaced with a circular slime mold of the same size sitting on oat. In the first and second treatments

(nutritive environments) we added glucose (100mM or 200mM) to the medium. In the third treatment (adverse environment), we added a known repellent (NaCl 100mM [94]) to the medium. Lastly, in the fourth treatment, the medium remained unchanged (neutral environment i.e. control treatment).

Subsequently, to investigate how chemotaxis modified the exploration behavior, we introduced discrete spots of attractants/repellents within a neutral substrate made of plain agar. In these so-called "spot experiments", we followed a procedure similar to that for the homogeneous environments. The arena consisted of a 14.5 cm diameter petri dish filled with plain 1% agar mixed with non-nutritive cellulose (5%). Once the agar had set, we punched two or three holes in a line configuration (diameter of each hole 2.5 cm). A circular slime mold (2.5 cm in diameter) was placed diametrically opposite to a glucose (attractant) spot of same size, placed 4.5cm away. In some of the treatments, a salt (NaCl 200mM, repellent) spot was added between the slime mold and the attractant spot. A total of 4 different treatments were tested: the first and the second with a single spot of glucose at concentrations of 100mM and 200mM respectively, the third and fourth keeping the glucose spots with the same concentrations and adding a NaCl 200mM spot. Previous experiments show no disruption of slime mold behavior due top punching alone [95].

All slime molds were fed just before the experiment so we assumed that they were in the same physiological state. The experiment consisted of a total of 8 different treatments. We replicated the experiment 20 times for each treatment and monitored each arena for 35 hours taking time-lapse photographs every 5 minutes. Figure 2.1 shows the experimental set-up for homogeneous environments (left) and discrete distributions of attractants/repellents (right).

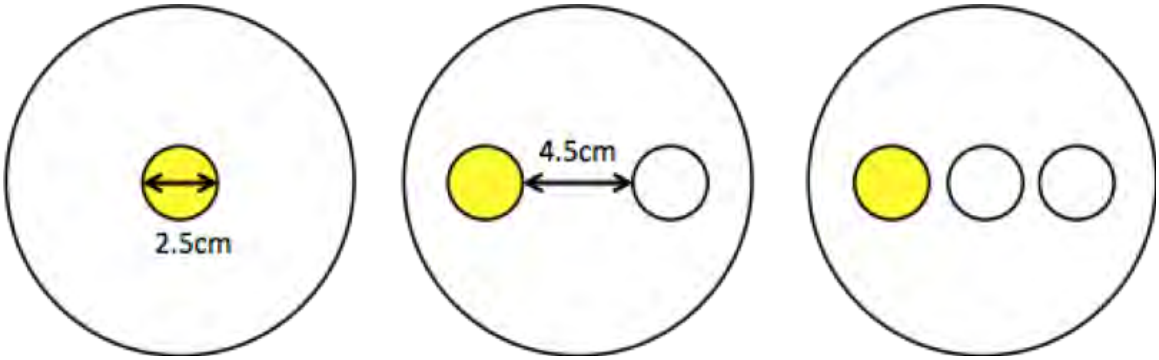

Figure 2.1: Experimental set-ups for homogenous environment and spot experiment.

2.3.4 Image Processing

Time-lapse images were taken every 5 minutes for a total of 420 pictures for each replicate. First, the outside of the petri dish was masked by manually identifying the circular contour; then, the images were converted into the ab* color space (which is the CLAB space without the L* lighting component) since this color space yielded better segmentation results. Then, the clustering algorithm k-means [96, 97] was used to classify every pixel into one of two categories or clusters: slime mold or not-slime mold. This last category was further refined based on pixel history, becoming either unexplored surface or mucus. Mucus is the substance left by slime after refinement, which acts as an external memory on explored areas [80].

The change of class from unexplored substrate to slime mold, defined as primary growth, means that slime mold reached a point it had never explored. Similarly, secondary growth is the change from mucus to slime mold, meaning that the cell is revisiting a location. Lastly, if the slime mold recedes, e.g. a pixel goes from slime mold to non-slime mold, it becomes mucus, and is defined as refinement. Note that once a pixel is identified as mucus, it can never be classified as unexplored substrate in the following time frames.

2.3.5 Image Analysis in Space and Time

We computed indexes to characterize slime mold geometry dynamics and averaged them over the 20 replicates to obtain statistically representative measures. In order to quantify the differences on distinct substrates, we first calculated the fraction of the petri dish area covered by slime mold, mucus and unexplored substrate over time. The total area, the lighting conditions and the test duration were the same for all treatments, both in the homogeneous and spot experiments. Note that glucose only provides energy to slime mold, which is not gaining significant mass during the experiments [95]. In other words, slime mold is changing its area by mostly by stretching and contracting, therefore changing its area density.

We then computed the cumulative area of primary growth, refinement, and secondary growth over the full period of the experiments comparing two consecutive images at the time. The cumulative area covered by primary growth is indicative of the total area of exploration, therefore it is always smaller or equal to the total area of the dish. The cumulative area covered by secondary growth indicates whether slime mold expansion is monotonic (dominated by primary growth) which results in a smaller magnitude, or cyclic (secondary growth dominated, with pulsatile movements) which results in a larger magnitude. The cumulative area covered by refinement indicates slime mold density changes. Within a given time interval, if the area covered by primary plus secondary growth equals that covered by refinement, then slime mold keeps the same density, whereas if it is superior, the slime mold stretches (e.g. density decreases). If secondary growth is negligible and if the area covered by primary growth equals the area covered by refinement, then slime mold displaces mass.

We calculated the extent of growth for each pair of consecutive images as the distance from each pixel where growth occurred (both primary and secondary) to the closest pixel classified as slime mold in the previous image. We calculated the migration rate as the ratio between the maximum extent of growth and the time interval between images (5 min). Then we delineated the region explored by slime mold within the interval as the contour of the

slime mold with an offset distance corresponding to the migration rate (see supplementary material in [27] for more details).

We estimated the fraction of secondary growth relative to the total number of pixels in the region of expansion. We then calculated the fraction of "expected secondary growth", which would have occurred if secondary growth had happened randomly. If the measured secondary growth fraction is higher (respectively, lower) than the expected one, this means that slime mold has a bias towards mucus (respectively, unexplored substrate).

Additionally, we computed four shape parameters indicative of the contour of slime mold: circularity, eccentricity, solidity and number of clusters. Circularity (C) is defined in the following equation:

$$C = \frac{P^2}{4\pi A} \tag{2.1}$$

Where P and A are the perimeter and area of the shape of slime mold at a given time. Eccentricity (E) is calculated as the ratio between the distance between the foci and the major axis length, as shown in the following equation:

$$E = \sqrt{1 - (b/a)^2} \tag{2.2}$$

In which a and b are the lengths of the major and minor axes, respectively. Solidity (S) is the ratio between the area of the slime mold contour and the area of its convex hull. Lastly, we measured the number of clusters by performing an erosion operation along the contour of slime mold, after which, only the clusters of high concentration of slime mold remained, which provided the number of pseudopodia or clusters of slime mold.

For the spot experiments, we also determined the distance from the slime mold to the glucose spot at every time, as the minimum distance between the contour of slime mold and the glucose spot. The evolution of the distance to glucose over time was analyzed in a way similar to a survival analysis.

2.3.6 Statistics

The full description of the statistics is provided in the supplementary material published in [27], which includes the statistical procedure and results and an R markdown allowing analysis reproduction. When dependent variables lasted until the occurrence of certain event, we conducted survival analyses using the R package coxme [98]. For the remaining dependent variables, we did linear analyses using the R packages lme4 [99] and lmerTest [100]. We tested as fixed factors the four treatments in the homogeneous experiment and both the attracting and the repelling spots and their interaction. The date of the experiment was considered as a random factor. Finally, we performed a nested model comparison using the R package MuMIn [101] by ordering models according to their Akaike criterion and represented the selected model by plotting each estimator, their 95%CI and p-values.

2.4 Results

2.4.1 Homogeneous environment

In order to study the influence of the environment on slime mold expansion rate, we analyzed the areas covered by slime mold, unexplored substrate and mucus over time, as shown in Figure 2.2.

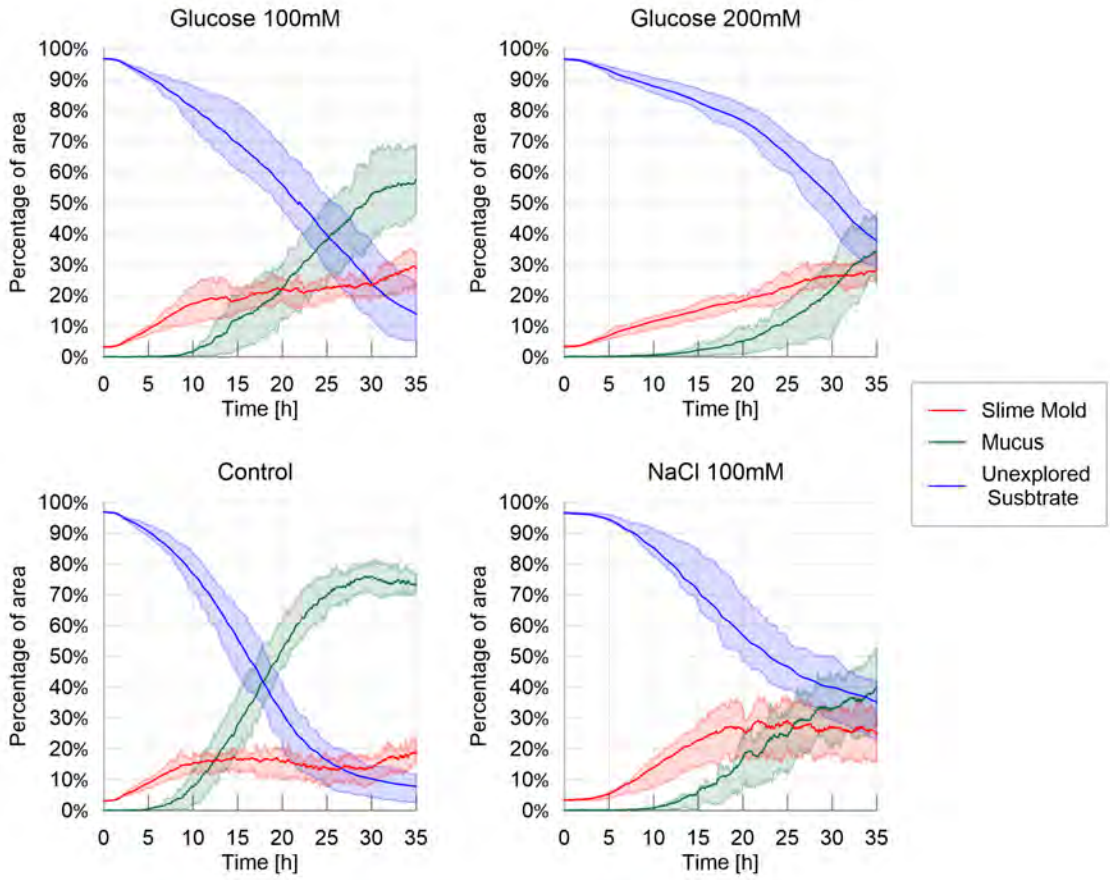

Figure 2.2: Fraction of area covered by slime mold, mucus and unexplored substrate – Homogeneous environment. The solid lines correspond to the average index calculated over the 20 replicates, while the shaded areas correspond to the first and third quartiles of the data, the dashed line correspond to the ratio between the average mucus area over the average slime mold area over time.

In a neutral (control) and slightly nutritive environment (glucose at 100 mM), the slime molds started to spread from the very beginning of the experiment (Figure 2.2; $P > 0.05$). By contrast, in a highly nutritive environment (200 mM glucose), slime molds started to explore later ($P < 0.001$ when compared to the control). Slime molds placed in an adverse environment (100 mM NaCl) were lagging the most and only started exploring after 3 hours ($P < 0.001$ when compared to the control). Once the slime molds started to explore, they all grew at the same rate (Figure 2.2; $P > 0.05$ when compared to the control) except the ones placed in a highly nutritive environment which were slowed down ($P < 0.001$ when

compared to the control). At the end of the experiment (after 35 hours), the slime molds reached a similar surface area in a control environment and in an adverse environment ($P > 0.05$ when compared to the control). Interestingly, after reaching a plateau at 18 hours, the area covered by the slime molds in an adverse environment oscillated with seemingly cyclic fluctuations (Figure 2.2). In both a slightly and a highly nutritive environment, the slime molds reached a higher final surface area than the slime molds placed in a control environment ($P < 0.001$ in both comparisons) and covered approximately 30% of arena at the end experiment. It is worth noting that in a highly nutritive environment, the area of the surface covered by the slime molds never reached a plateau after 35 hours, suggesting that the slime molds did not reach its maximum surface area (Figure 2.2).

Refinement i.e. appearance of mucus, was observed after 5 hours in the control environment. In all other environments, mucus appeared later ($P < 0.001$ for all treatments when compared to the control). In a highly nutritive environment, mucus was only observed after 10 hours, which marked the strongest delay in the refinement process. Once the mucus started to be apparent, its surface area grew quicker in the control environment than in the other three treatments ($P < 0.001$ for all treatments when compared to the control). Thus the area of the surface covered by mucus at the end of the experiment was the largest in the control environment where it reached 75% of the arena against 55%, 40% and 35% for the slightly nutritive, the adverse and the highly nutritive environments respectively ($P < 0.001$ for all treatments when compared to the control).

Hence, slime molds placed in a control environment explored almost all the arena leaving only 5% of the arena unexplored while in the other treatments the area of the surface unexplored were significant: 15%, 35% and 38% for the slightly nutritive, the highly nutritive and the adverse environments respectively. Interestingly, although the growth rate dynamics differed between highly nutritive and adverse environments, the final unexplored surface areas were similar. In a highly nutritive environment the slime molds grew slowly and steadily while in an adverse environment slime molds grew rather quickly but after a

long delay.

Next, we analyzed the evolution of the cumulative areas covered by primary growth, refinement and secondary growth (Figure 2.3). The cumulative area covered by secondary growth, which reveals the cyclic nature of the exploration process, was the highest in the adverse environment (480% coverage) followed by the control environment (380%), the slightly nutritive environment (250%) and the highly nutritive environment (180%). All comparisons lead to significant differences $P < 0.05$, except control vs. adverse environment. This observation confirms that exploration was slowed down by the presence of nutrients, and that the pulsatile behavior (i.e the exploration of previously explored area) was stimulated by repellents.

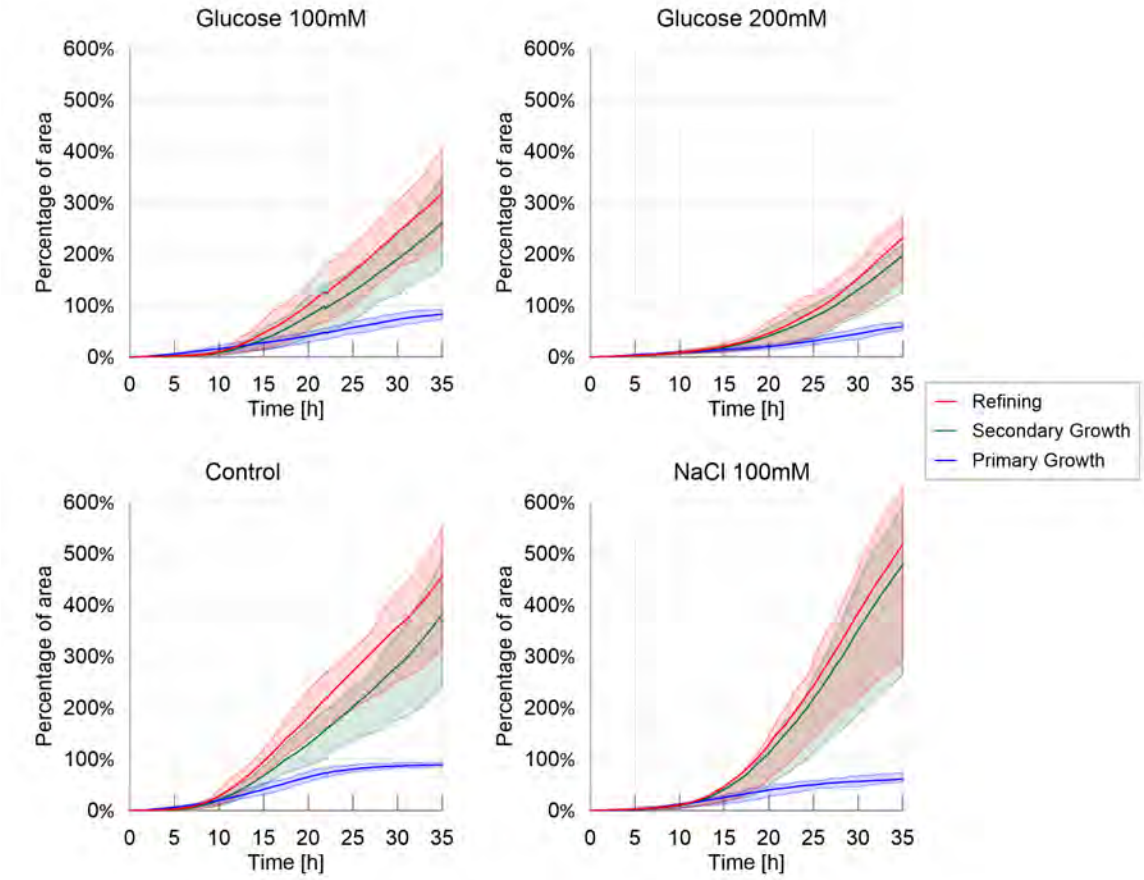

Figure 2.3: Cumulative areas covered by primary growth, refinement and secondary growth – Homogeneous experiments. The solid line corresponds to the average index calculated over the 20 replicates, while the shaded areas correspond to the first and third quartiles of the data.

In accordance with the previous results, Figure 2.4 (average migration rate for the treatments) shows a higher migration rate for the control treatment than for the other treatments ($P < 0.001$ for each pairwise comparison). While slime molds exploring the highly nutritive environment were slower than slime molds exploring the slightly nutritive or the adverse environment ($P < 0.001$ each), these two showed no significant differences ($P > 0.05$).

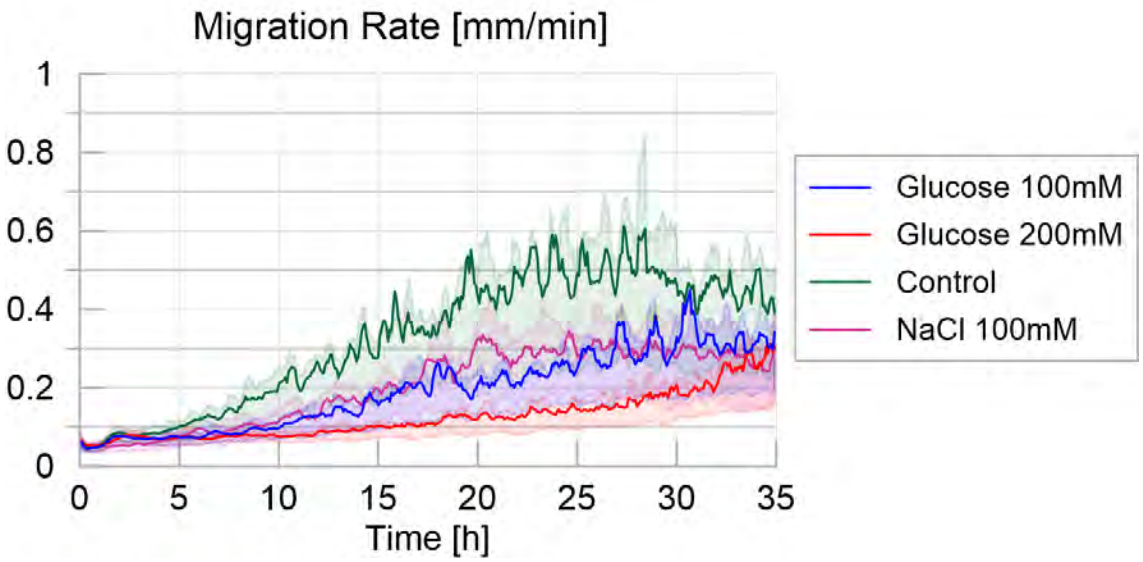

Figure 2.4: Migration rate over time for the four different treatments, defined as the maximum distance between the contours of the slime mold between two consecutive images divided by their time interval (5 minutes apart), measured in millimeters per minute. The solid line corresponds to the average calculated over 20 replicates per treatment, while the shaded areas correspond to the first and third quartiles of the data.

The slime molds exploring the adverse environment showed the highest probability to explore a previously explored substrate than the other treatments as shown in Figure 2.5 ($P < 0.001$ for each pairwise comparison). When exploring a highly nutritive environment, slime molds also displayed a significant positive tendency for secondary growth ($P < 0.001$ for each pairwise comparison) but significantly less strong than on the adverse environment ($P < 0.001$). For the other treatments, the measured proportion of secondary growth was not different from the expected proportion of secondary growth, indicating that slime molds did not avoid previously explored substrate and explored randomly. The peaks observed within the first 5 hours of the experiment correspond to an isotropic extension immediately followed by a refinement process that occurred before the slime mold started to explore continuously its environment. This behavior is often observed when a slime mold is introduced in a new environment and is referred as "contemplative" [102] i.e. the slime mold migrates, retracts and moves again. The peak was larger in an adverse environment.

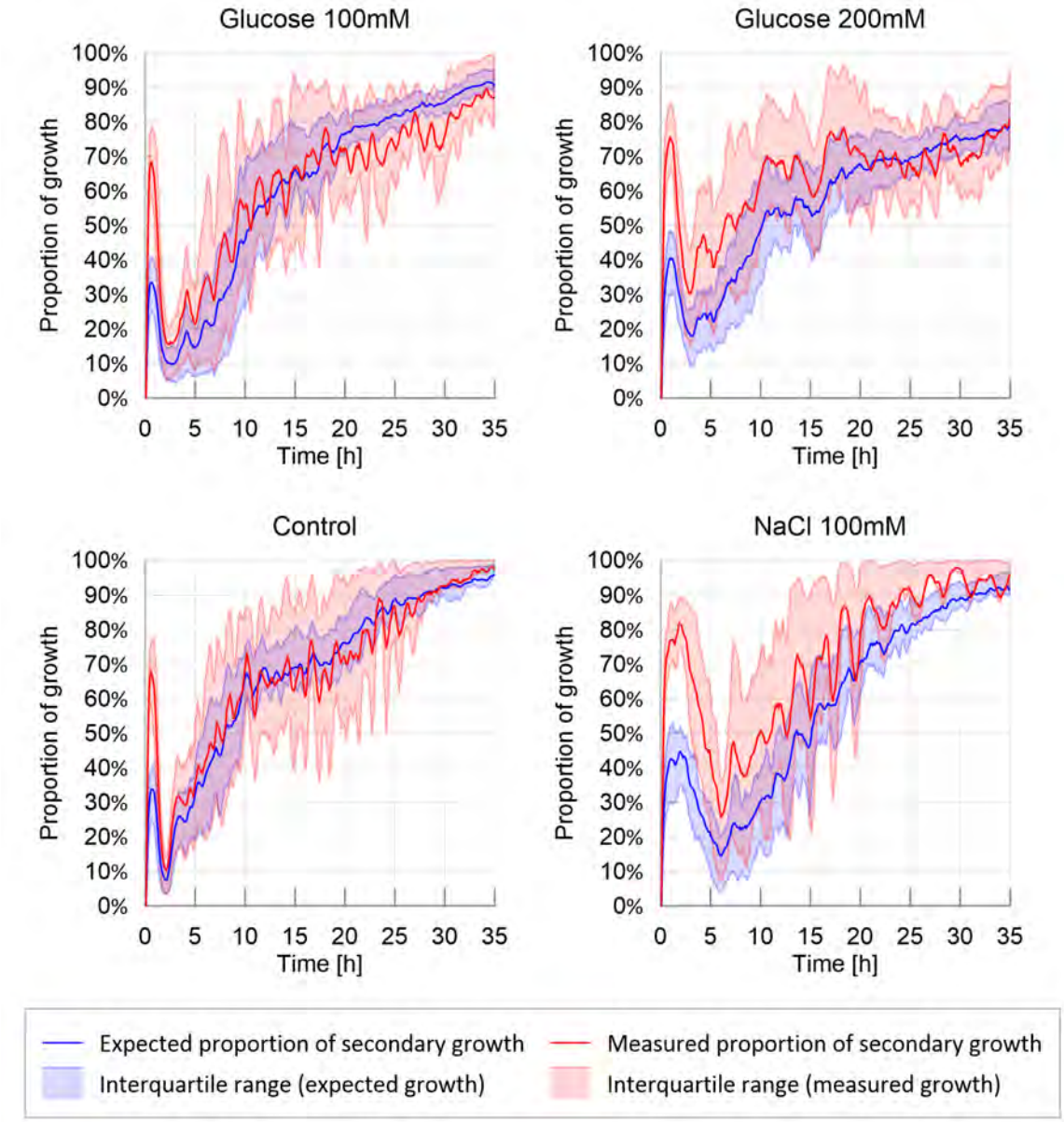

Figure 2.5: Ratio of secondary growth: observed and expected proportion of secondary growth. The solid line corresponds to the average calculated over 20 replicates per treatment, while the shaded areas correspond to the first and third quartiles of the data.

We then analyzed the evolution of the shape of the slime molds contour. Note that the experimental set ups in which slime molds were placed exhibited radial symmetry. Hence, no preferential expansion direction was expected. We thus focused on contour shape, not orientation. As expected, circularity was initially one in all tests (circular slime mold spot),

and increased over time as the contour shape departed from a circle. In the control and nutritive environments, circularity remained between 1.05 and 1.10, whereas it fluctuated between 1.05 and 1.30 in the adverse environment. This observation suggests that, in an adverse environment, slime molds explored the petri dish by spreading and thinning over larger areas than in the other environments, which led to shape changes and a decrease of slime mold circularity. However fluctuations among the 20 replicates were too high to identify any trend in the evolution of slime mold circularity.

The eccentricity index was initially close to zero (circular cell) and increased up to almost 0.8 over time, with important fluctuations in all the treatments. Eccentricity is an indicator of the number of pseudopodia. But a non-eccentric convex hull can enclose non-circular contours of slime mold, since pseudopodia can develop in a symmetric fashion. That is why no major difference was noted between the treatments. This result highlights the absence of preferred expansion direction in symmetric, homogeneous environments. Solidity decreased with the emergence of pseudopodia, since slime mold branching disrupted the initially convex shape of the slime mold (Figure 2.5). In the control environment, in which the exploration rate was the highest, the decrease of solidity of the slime mold area was the highest (and the fastest) decreasing from 1 to 0.3 and then becoming relatively stable, with fluctuations of +/- 0.05 ($P < 0.01$ for all paired comparisons except adverse environment vs. slightly nutritive environment, where $P > 0.05$). Slower exploration resulted in a slower and steadier loss of solidity as observed in the nutritive and adverse environments. The highly nutritive environment yielded the highest solidity at the end of the experiment (0.6), which confirmed that the presence of glucose slowed down exploration.

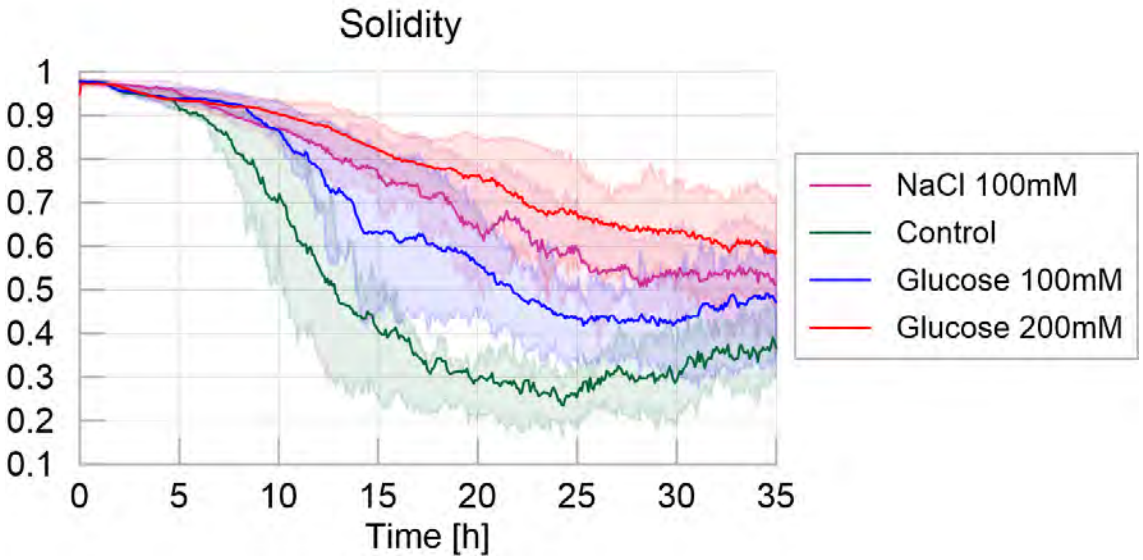

Figure 2.6: Solidity - Homogeneous experiments. The solid line corresponds to the average index calculated over the 20 replicates, while the shaded areas correspond to the first and third quartiles of the data.

We next focused on the number of clusters, corresponding to the number of pseudopodia (Figure 2.7). Initially the slime molds stretched as a single cluster. Once mucus started to be apparent, slime molds usually divided up into several clusters, and started the active exploration phase. In highly nutritive and adverse environments, the number of clusters over time was lower than in the other two treatments ($P < 0.01$ for all paired comparisons except for adverse environment vs. highly nutritive environment). This observation confirmed that the presence of concentrated nutrients slowed down the exploration, and that the presence of repellents triggered a highly pulsatile behavior with small exploration fronts, which were sometimes not detected as separate clusters.

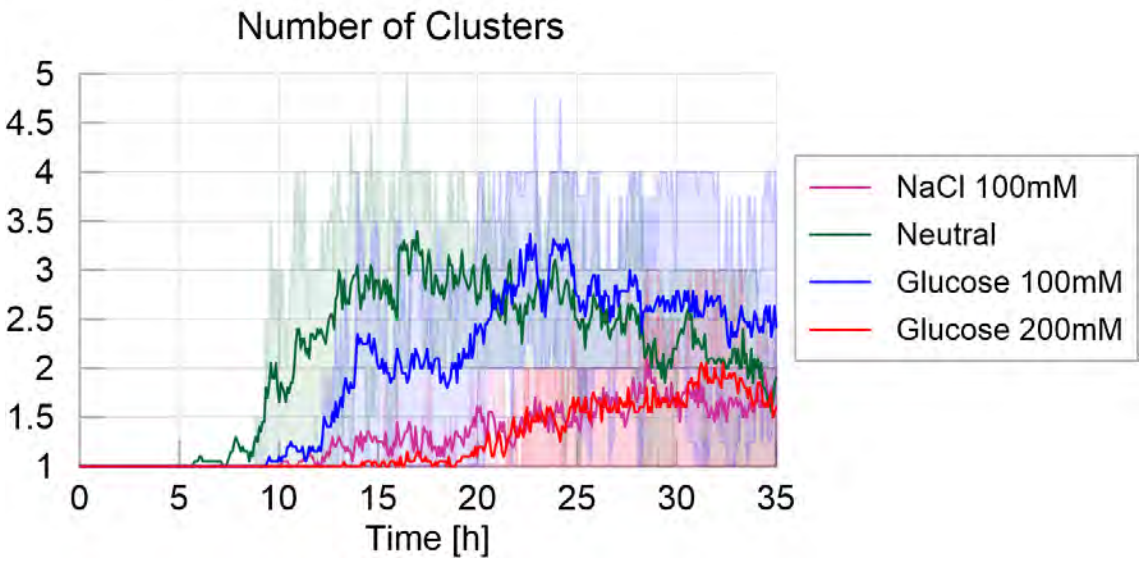

Figure 2.7: Number of clusters - Homogeneous experiments. The solid line corresponds to the average index calculated over the 20 replicates, while the shaded areas correspond to the first and third quartiles of the data. Pictures show examples of clusters for the Control environment (left) and 200mM Glucose environment (right).

2.4.2 Spot experiments

In the spot experiments, we studied the influence of discrete distributions of nutrients and repellents on exploration dynamics. When looking at the evolution of slime mold, mucus and unexplored substrate over time (Figure 2.8), we only observed marginal difference among the treatments, which all exhibited similar patterns of exploration, e.g. similar percentage of non-explored area and similar mucus accumulation. The presence of an adverse spot only delayed the appearance of the first pseudopod ($P < 0.05$) but not the first appearance of mucus. The only noticeable differences lie in the surface area reached at the end of the experiment: slime molds that were offered a highly nutritive spot grew larger ($P < 0.01$). By contrast, the area of the surface covered by mucus was lower (mucus final surface:$P < 0.01$) than slime molds that were offered a slightly nutritive spot. In comparison with the experiments conducted in homogeneous environments, we did not observe any expansion/refinement cycles in the spot experiments, meaning that slime mold spread steadily towards the food source despite the presence of an obstacle on the way.

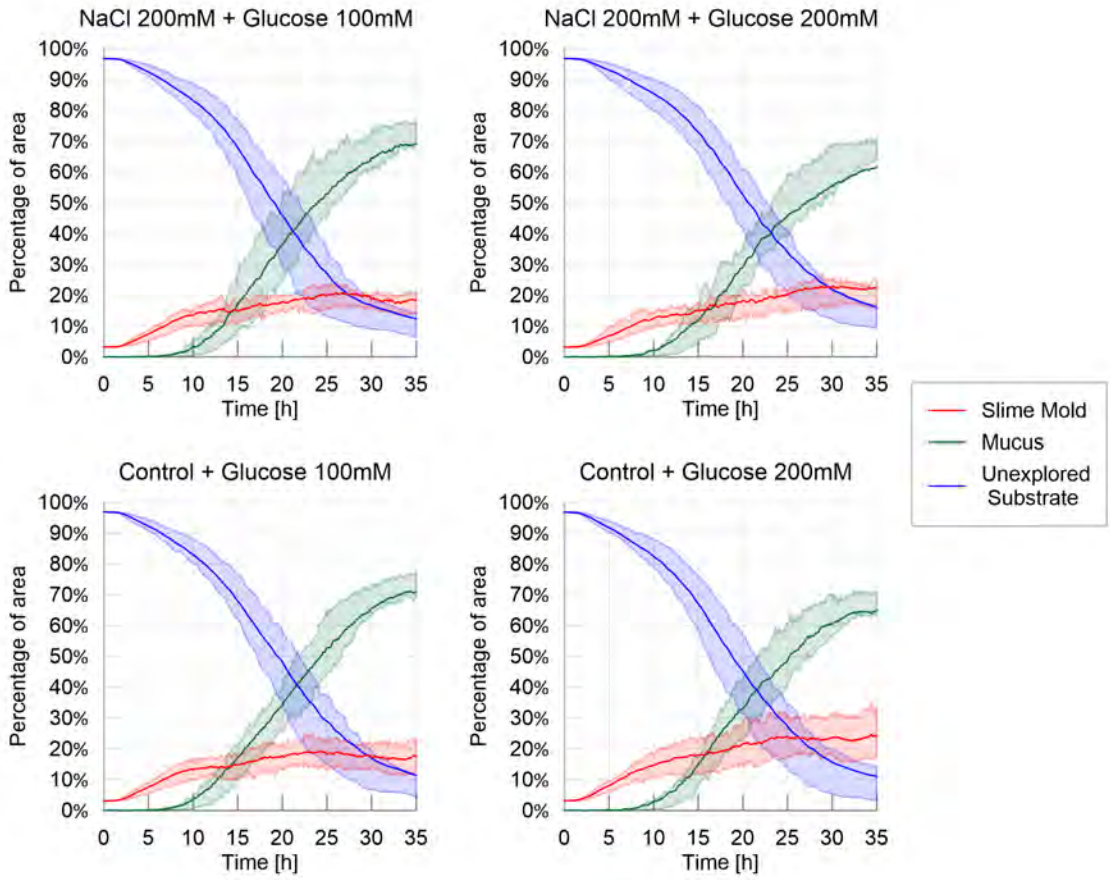

Figure 2.8: Fraction of area covered by each entity (slime mold, mucus, unexplored substrate) – Spot experiments. The solid line corresponds to the average index calculated over the 20 replicates, while the shaded areas correspond to the first and third quartiles of the data.

The exploration behavior in the spot experiments was similar to that observed in the control environment in homogeneous experiments as observed in Figure 2.9, which includes the average percentage of unexplored area for the homogeneous and spot experiments shown in Figure 2.2 and Figure 2.8 respectively. This suggests that the spatial exploration of slime mold depended mostly on the substrate and not on the geometric distribution of the nutritive and adverse stimuli.

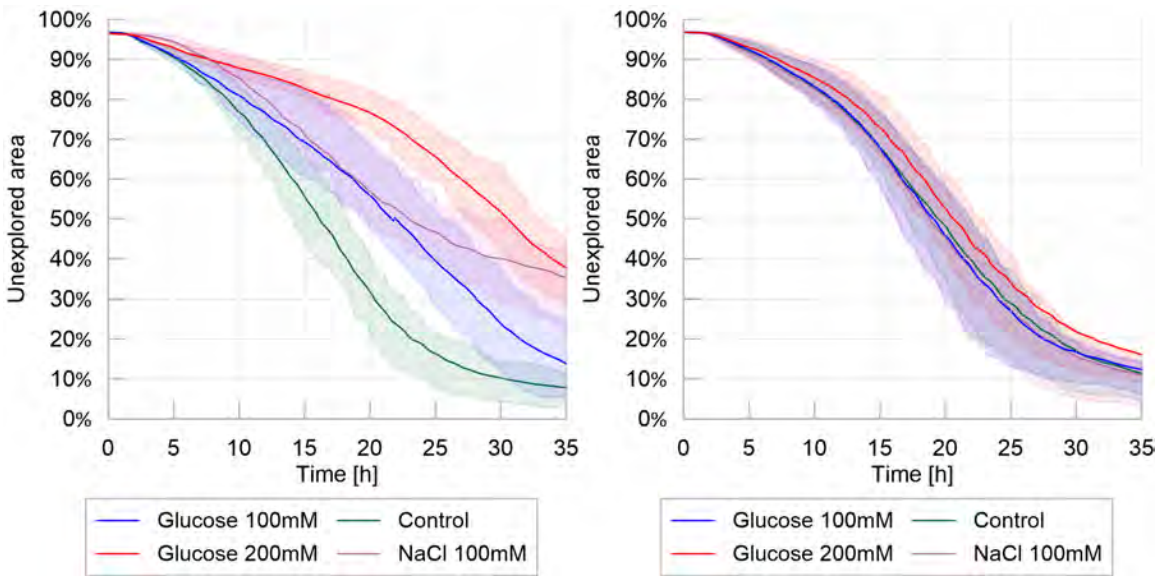

Figure 2.9: Exploration behavior: homogeneous experiments (left) vs. spot experiments (right). Percentage of unexplored area over time. Mean values of over 20 replicates for each different treatment.

The cumulative areas covered by secondary growth for the spot experiments (Figure 2.10) were also similar for all treatments ($P > 0.05$), suggesting again, that isolated spots with nutritive or adverse stimuli did not alter the overall exploration of slime molds when growing on the same, control, substrate.

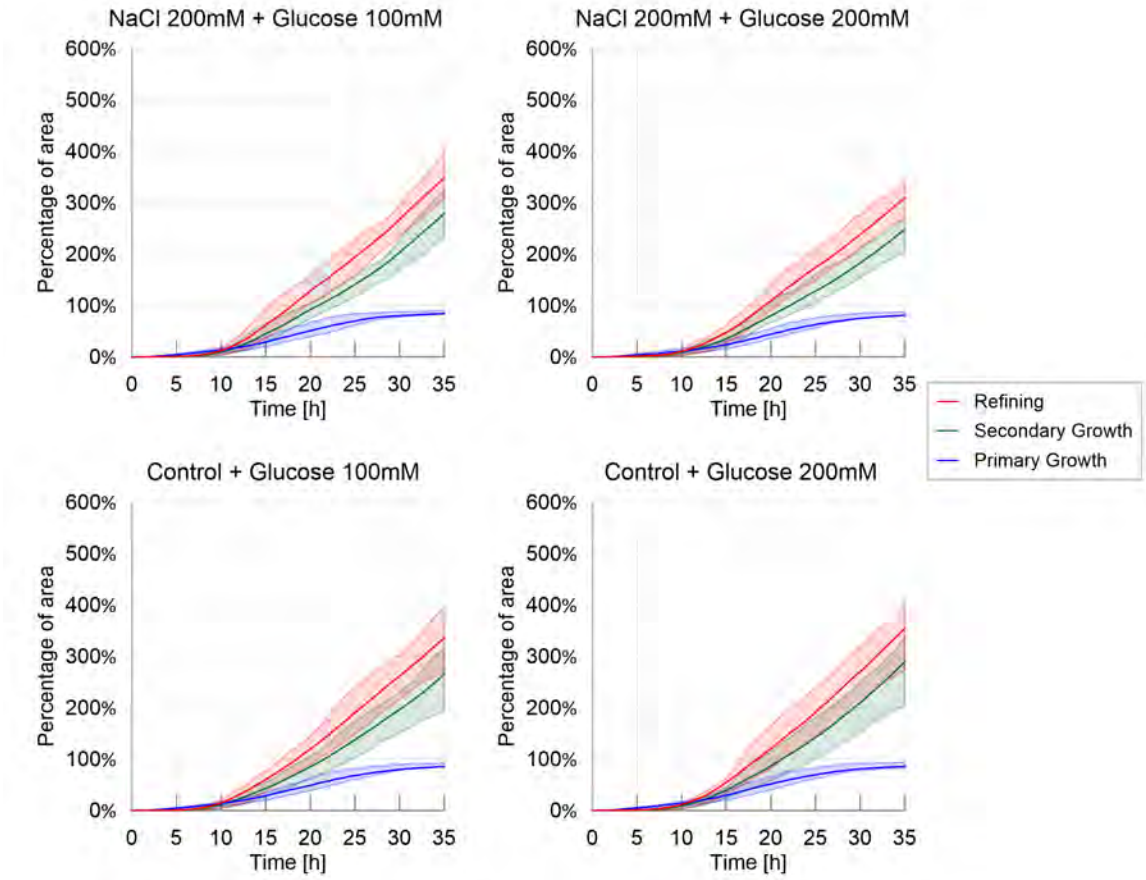

Figure 2.10: Cumulative areas covered by primary growth, refinement and secondary growth – Spot experiments. The solid line corresponds to the average index calculated over the 20 replicates, while the shaded areas correspond to the first and third quartiles of the data.

The trends of migration rate, as shown in Figure 2.11, show that slime molds were not affected by the difference between treatments, showing only a slight effect of the concentration of the food spot ($P < 0.05$). This effect showed that the migration rate was slightly superior when slime molds were offered a higher than a lower nutritive spot.

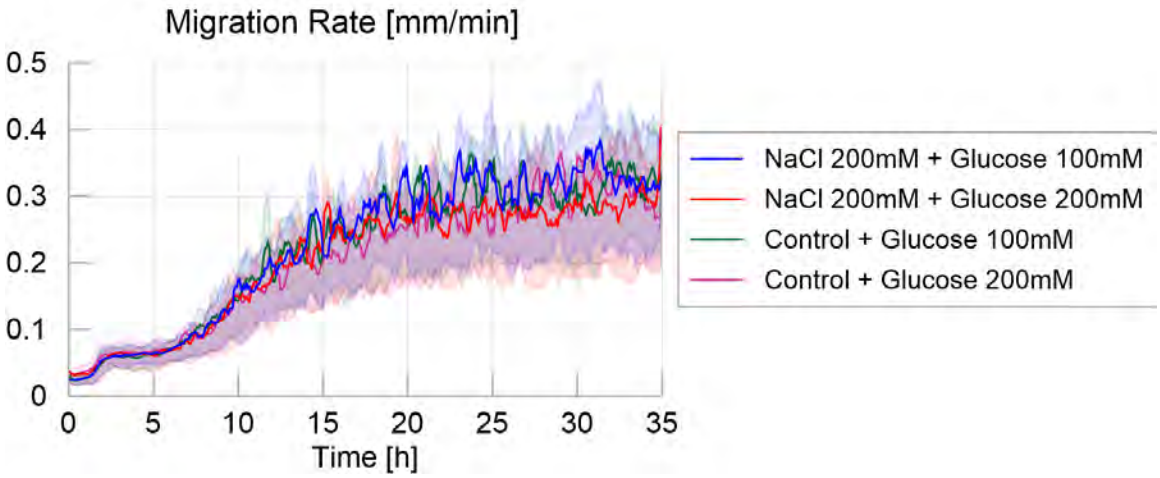

Figure 2.11: Migration rate over time for the four different treatments, defined as the maximum distance between the contours of the slime mold between two consecutive images, divided by their time interval (5 minutes apart), measured in millimeters per minute. The solid line corresponds to the average calculated over 20 replicates per treatment, while the shaded areas correspond to the first and third quartiles of the data.

Similarly, looking at the predilection of slime molds to grow towards mucus (secondary growth), as shown in Figure 2.12, no significant differences were observed between treatments, which suggests that the growth type is not influenced by the existence or concentration of discrete attracting/repelling spots.

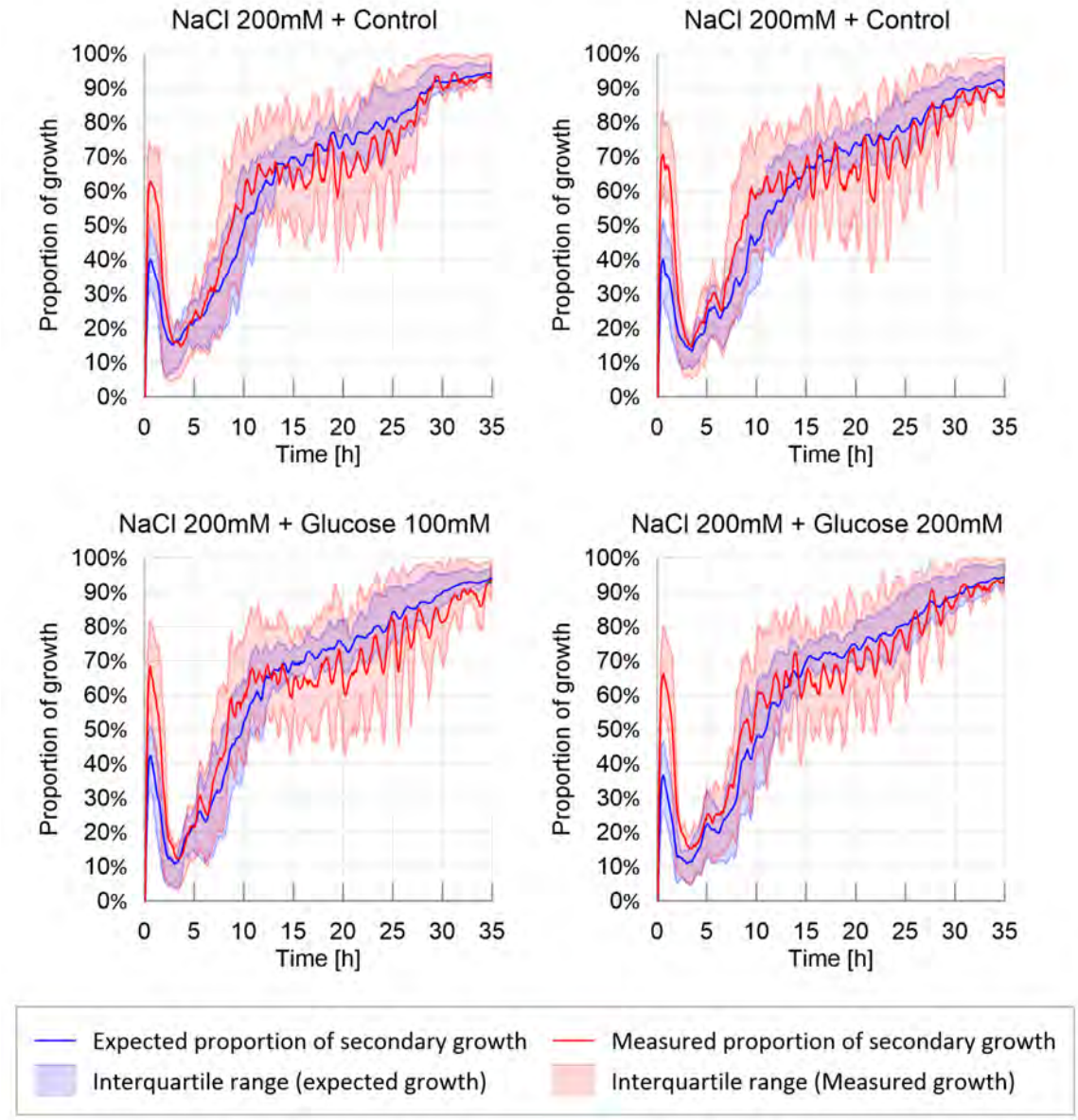

Figure 2.12: Probability of secondary growth: observed and expected proportion of secondary growth. The solid line corresponds to the average calculated over 20 replicates per treatment, while the shaded areas correspond to the first and third quartiles of the data.

The four different shape indexes for the spot experiments support the hypothesis that discrete spots of nutrients or repellents did not affect the overall expansion dynamics and exploration cycles. This interpretation is confirmed by the average number of pseudopods, which was found to be correlated to the formation of mucus during the exploration phase:

in all the spot experiments, the number of clusters increases from 1 to 2.5 within around 12 hours, to reach a plateau afterwards. In other words, less exploration cycles were observed in non-homogenous environments, yielding less pseudopodia.

The evolution of the shortest distance from the slime mold cell to the glucose spot is shown in Figure 2.13, which can be viewed as a "survival" plot, displaying the proportion of the replicate (P) in which slime mold has not reached the nutritive spot at a given time. Both increasing the concentration of nutrient in the spot (from slightly nutritive to highly nutritive) and adding an adverse spot increased the time to reach the food patch by decreasing the probability to reach it (time to food patch: nutritive: $P < 0.01$; adverse: $P < 0.05$).

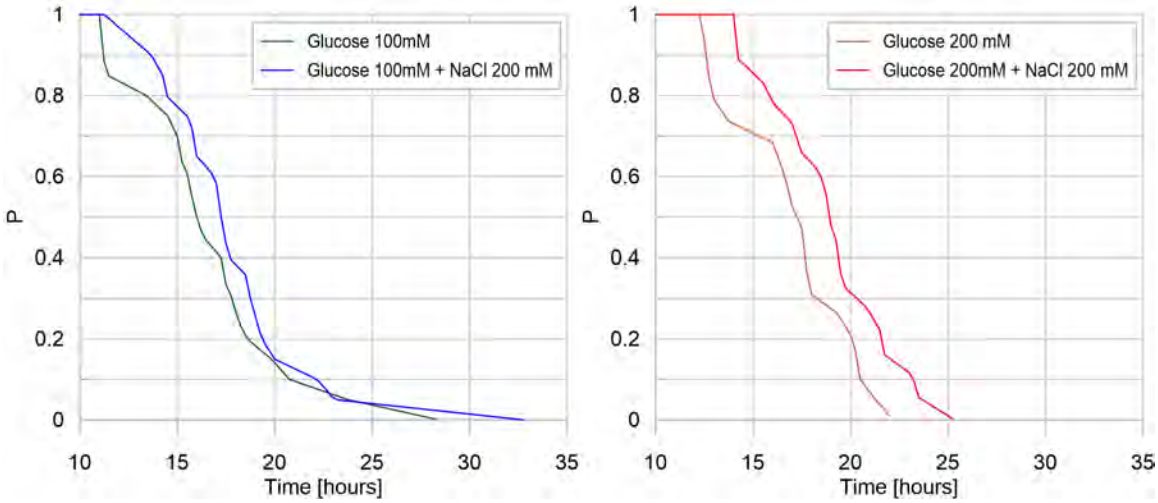

Figure 2.13: Survival plot: glucose concentration effect. For each treatment, on the vertical axis, the value P corresponds to the fraction of replicates that have not reached the glucose spot at a given time. For a representative number of replicates, P is the probability that glucose has not been reached by slime mold at a given time, for a specific treatment. On the horizontal axis, each value of time corresponds to the average time it takes for a certain fraction of slime mold (P) to reach the glucose spot.

2.5 Discussion

Exploration in slime molds involves two different processes: area extension and movement [72]. Slime molds locomotion and morphogenesis depend on the response of the organism

to the environmental conditions such as light, hygrometry, pH, or the presence of chemicals [73, 103]. In the present paper, we show that distributions of nutritive and adverse cues affected drastically the exploration pattern of slime molds.

The typical exploration behavior of a slime mold (in control conditions) started by a stretching period where the slime molds grew uniformly in all directions for 10 hours. Then, the contour of slime mold lost circularity when the first pseudopodia appeared, which also corresponds to the first occurrence of mucus. This phenomenon is typical of the directed digitated growth, or branching phase, described by numerous authors [104, 105, 88, 94, 72]. Slime molds developed multiple pseudopodia and did not exhibit any preferential exploration orientation. At the end of the experiment, almost all the arena was explored by the slime mold. Thus the exploration was characterized by three phases: (i) Primary growth only, in the quasi-absence of mucus (5-10 hours); (ii) Combination of primary and secondary growth; (iii) Secondary growth only, when the slime mold stops exploring new substrate areas.

Regarding the result obtained for the homogeneous distribution of nutritive cues, we noted first, an environment containing a uniform distribution of nutrients slowed down the exploration of slime mold, mainly by delaying secondary growth and increasing the period of the pulsatile exploration/refinement movements. The area not explored by the slime molds was 3 times larger (respectively 7 times larger) than in the environment deprived of nutrient (control case) for a slightly nutritive (respectively highly nutritive) environment. The exploration rate was almost linear for highly nutritive environments, while for other treatments, the area covered by the slime molds reached a plateau after a period of stretching, which indicates secondary growth and slime mold displacement. This means that slime molds that explored nutritive environments never exhibited a Phase (iii) in their exploration pattern.

Second, on substrates with higher nutrient concentrations, the slime molds grew in a more compact fashion, i.e. slime molds presented the highest solidity index and the lowest

number of pseudopodia (clusters). Additionally, the appearance of mucus, which indicates that the slime mold was withdrawing, occurred much later in nutritive environments. As glucose is only adverse when above 300mM, our results suggest that nutritive media depressed migration due to feeding behavior. This allows the organism to remain at a site until nutrients are exhausted [106, 95, 107]. In previous studies, it was shown that the surface area of substrate covered increases when slime mold responds to nutrient dilution [95, 108]. Here, we confirmed these observations and noted that slime mold tended to migrate and grow faster on substrates with the lowest concentration of nutrients, thus maximizing nutrient intake and optimizing its trade-off with foraging.

Similarly, for a homogeneous distribution of adverse cues we noted the aversion of slime mold to salt manifested itself through longer contemplative behavior, delayed primary growth and a higher probability to crawl on previously explored surface. In addition, the slime molds grew more compactly and with less pseudopods. This suggests that slime molds were actively avoiding contact with the adverse surface [109, 85, 102]. In the absence of cell walls, slime mold has no other protection from the environment than mucus. Indeed, in bacteria for instance, mucus is used as protective barrier for the cells against harsh external conditions [110]. In slime molds, the extracellular mucus excreted by the slime molds can have different roles: hydrophilic shield to prevent water loss [111], a lubricating surface over which the slime molds can easily crawl [112], a defensive coat to protect against invasion by foreign materials and organisms [113, 111], an aid to phagocytosis [114], a surface that promotes ion-exchange [111] and has externalized spatial memory that helps navigation in unknown environments [115, 81, 70]. Here, we can add a new function for the mucus i.e. a buffer to move in adverse environment. In other words we demonstrate for the first time that the mucus may be used as a safe haven to avoid prolonged and repeated contacts with hazardous substances encountered in the environment.

Lastly, for non-homogeneous distribution of adverse and non-adverse cues, e.g. spot experiments, pulsatile movements were limited and slime molds responded in the same way

regardless of the concentration of glucose used as attractant and regardless of the presence of a salt spot on the way to the glucose spot. Slime molds grew in a more compact fashion, i.e. slime molds presented the highest solidity index and the lowest number of pseudopodia (clusters). Additionally, the evolution of the areas covered by slime mold and mucus over time indicates that the response of slime molds to heterogeneous environments was similar to that in the control case. This result suggests that in the spot experiments, the exploration behavior of slime mold is mostly controlled by the substrate. Our observation confirms that salt reception can be affected by the presence of sugars [91]. The authors in [91] showed that the "apparent" enthalpy change accompanying salt perception decreases with increase of sugar concentration.

The proposed image analysis program allows extracting information on expansion rates, geometric changes and probability of occupancy. Ongoing developments aim to acquire high quality images of slime mold exploration tests and to expand the code's capabilities for extracting topological information on the networks formed by slime molds. Slime molds have been shown to be capable of distributed sensing, parallel information processing, and decentralized optimization [116, 103, 83]. It is also viewed by some researchers as an inspiration for bio-computing devices [117, 118, 119, 120, 121]. Our image analysis methods will help such research avenues by improving data collection.

CHAPTER 3
INFLUENCE OF SUBSTRATE IN CELL FUSION AND NETWORK DYNAMICS OF SLIME MOLD

3.1 Aim and summary

The present chapter follows up on the study of the behavior of slime mold (*Physarum polycephalum*), specifically focusing on networks dynamics. Slime mold cells are remodelled constantly in response to mass gain/loss, environmental conditions and fusion between adjacent cells. Understanding how slime molds networks are built and fuse to allow for efficient exploration and adaptation to environmental conditions, is currently an open area of investigation. Here, we characterize the network organization of slime molds exploring homogeneous neutral, nutritive and adverse environments. The contents of this chapter has been submitted for publication and is under review at the time when this thesis is written [28].

We developed a fully automated image analysis method to extract the network topology and we followed the slime molds before and after fusion. Our results show that: (1) slime molds build sparse networks with thin veins in a neutral environment and more compact networks with thicker veins in a nutritive or adverse environment; (2) slime molds construct long, efficient and resilient networks in neutral and adverse environments, whereas in nutritive environments, they build shorter and more centralized networks; and (3) slime molds fuse rapidly and establish multiple connections with their clone-mates in a neutral environment, whereas they display a late fusion with fewer connections in an adverse environment. Our study demonstrates that slime mold networks evolve continuously via pruning and reinforcement, adapting to different environmental conditions.

3.2 Introduction

Transportation networks where fluids are transported from one point of the network to another are ubiquitous in nature. Vascular networks in animals, plants, fungi and slime molds are commonly cited examples of such natural transportation networks. These networks are often studied as static architectures, although most of them have the ability to alter their morphology in space and time in response to environmental conditions [122]. Morphological alterations often include short term changes in the vein diameter and long-term structural adaptation such as addition or loss of veins [122, 123]. The slime mold *Physarum polycephalum* is a unicellular organism that is often used to study problem-solving in single-celled organisms [74, 124, 125, 89, 126, 127] and constitutes an ideal model system to study short and long term alterations in network morphology, in response to environmental conditions.

The motion and behaviour of *Physarum polycephalum* rely on a complex internal architecture which consists of a complex network of interconnected veins. The main function of this vein network is to transport oxygen and nutrients to maintain homeostasis in cells that range from 10 square micrometres to 10 square meters [128]. These veins contract and relax periodically, causing the cytoplasm to flow back and forth – a phenomenon called "shuttle streaming". These contractions produce a pressure gradient that pushes the cytoplasm towards the cell periphery where the veins vanish, forming fan-like pseudopods. The slime mold membrane extends and retracts in synchrony with the shuttle streaming, allowing cell migration at a speed up to few centimeters per hour [129]. In a homogeneous environment, the exploration process of slime mold alternates between refinement and generation of pseudopods that explore the domain [27].

The frequency and the amplitude of the vein contractions depend on external cues [130]. For instance, when *Physarum polycephalum* perceives a localized chemical attractant such as glucose in the environment, the veins contract at a high rate and the membrane migrates

toward the attractant [107, 131, 132], whereas when it senses a localized chemical repellent such as Sodium Chloride (NaCl), the veins contract at a lower rate and the membrane retracts, moving away from the repellent [76]. It has been shown that slime molds growing in an environment where NaCl is homogeneously distributed slow down the exploration process to allow production of mucus protection, whereas slime molds that grow in an environment where glucose is homogeneously distributed slow down the exploration process to allow metabolization [27]. Hence, *Physarum polycephalum* can adjust its behaviour, shape, size and migration speed based on environmental conditions. Interestingly, as the slime mold moves and expands, its networks of veins are also evolving and growing [133]. A question that arises is how external cues affect network morphology.

Physarum polycephalum cells have been modeled as undirected graphs by a number of authors, which even resulted in the creation of a public repository of slime mold extracted graphs [134], and different codes for network extraction from raw images [135, 136], which describe different segmentation networks based on the characteristics of the acquired images. Different studies on the network dynamics of slime molds suggest that the networks are hierarchical, with the distribution of veins widths and lengths following exponential, gamma or log-normal distributions, with the majority of the veins being short and thin, with a few long and thick veins [137, 138, 139]. As is the case in biological networks, the mean node degree is close to three in *Physarum polycephalum* networks, with little variations due to the existence of end branches. Thicker and longer veins have higher centrality than their adjacent veins [140, 136]. Smaller, less central veins allow adaptability and redundancy of the network as it constantly evolves and coarsens as a result of exploration [139, 136]. Even though the majority of studies on the network dynamics of *Physarum polycephalum* have focused on neutral environments, a recent study [141] revealed that changes in vein diameters can be induced by localized chemical attractants. However, the impact of these changes on the complete network morphology was not investigated. Ito et. al. [138] studied the effect of attractive and mildly repulsive environments on the network

formed by *Physarum polycephalum*, showing that nutritive environments usually yield networks with thinner veins but longer overall networks than adverse environments, which promote thicker veins but shorter networks.

Another interesting feature of *Physarum polycephalum* is that it can be severed into viable and structurally similar yet smaller cells. Upon contact, these cells can fuse with each other. Slime mold fusion constitutes a defining feature of the lifestyle of slime molds and allows them to share information [94]. When two different networks merge, remodeling might involve changes in vein diameter as well as pruning of supernumerary veins as seen in animal vascular systems [142]. But to our knowledge, the reorganization of *Physarum polycephalum* networks after fusion has never been studied.

Hence, due to its extremely original behaviour, fast migration rate and outstanding network topology, the acellular slime mold *Physarum polycephalum* offers an attractive model for the analysis of morphogenesis dynamics underlying cellular migration, exploration and fusion. In this chpater, we grow two slime mold cells of the same strain in adverse, nutritive and neutral environments in order to characterize the morphology, network and dynamics of *Physarum polycephalum* before and after fusion. First, we investigate whether and how the topology of slime mold networks is affected by different environmental conditions, by comparing an adverse environment (using sodium chloride NaCl as a repellent), a nutritive environment (using glucose as a chemo-attractant) and a neutral environment (using plain agar). Second, we analyze the evolution of slime mold networks within the three hours after fusion. We develop a program that automatically analyzes sequences of images to track the area and shape of the surface covered and explored by the slime mold, transforms images of slime mold into undirected graphs, and calculates network cost, efficiency and resiliency indexes. Contrary to tools proposed in previous studies [138, 135, 136], our program is flexible on the image quality requirements, which allowed us to run multiple experiments simultaneously and to compute the cost and efficiency of complete networks.

3.3 Methods

3.3.1 Species

The slime mold *Physarum polycephalum* is a unicellular organism that belongs to the Amoebozoa group. Its vegetative state, the plasmodium, is a giant mobile cell that consists of a syncytium of nuclei and an intracellular cytoskeleton, which forms a complex cytoplasmic network of veins. Its cytoplasm consists of a viscous phase (ectoplasm) and a liquid phase (endoplasm) characterized by different concentrations of fibrous proteins. The ectoplasm, which contains actin and myosin, forms the contractile walls of the veins. Within these veins flows the endoplasm, which contains organelles such as nuclei and mitochondria. Ecto- and endo-plasms are convertible into one another. A starving plasmodium can enter a dormant stage, called sclerotium, and turn back to a plasmodium after being transferred to a fresh food medium. In this paper, we used the Australian strain provided by Southern Biological, Victoria, Australia. We revived a total of 12 sclerotia to conduct the whole study.

3.3.2 Rearing Conditions

The slime molds were reared on a 1% agar medium with rolled oat flakes. They were fed every day and their agar medium was replaced daily. Slime molds were 2 weeks old when the experiment started. All experiments were carried out in the dark, at a temperature of 25 degrees Celsius, a humidity of 80%, for 48 h. Pictures were taken with a Canon 70D digital camera.

3.3.3 Experimental Setup

To investigate how the substrate affects the vein network evolution before and after fusion, we observed the behaviour of two slime molds that explore a medium that is neutral, slightly nutritive or adverse. Two circular slime molds (13 mm diameter) were placed di-

rectly on two opposite sides of a circular arena that consisted of a 90 mm diameter petri dish filled with either a plain 1% agar (neutral environment, i.e. control treatment), a plain 1% agar mixed with glucose (100 mM, nutritive environment), or a plain 1% agar mixed with salt (100 mM, adverse environment). All slime molds were fed just before the experiment, so we assumed that they were in the same physiological state. We replicated the experiment at least 40 times for each substrate. We tested and monitored each arena for 48 hours taking time-lapse photographs every minute. An LED panel underneath the petri dishes was turned on for 3 seconds when photographs were taken using a self programmed Arduino connected to the camera.

3.3.4 Image Segmentation

The image analysis algorithm was implemented in Matlab [143]. First, we masked manually the area outside of the contour of the petri dish. Then, a Laplacian filter was applied to enhance the contrast of the images. A traditional segmentation using the RGB space did not yield good results because the color region of slime mold (light yellow) was too similar to the color values of the background (white). Therefore, we proposed a segmentation method based on the LAB color space. Specifically, we used the components L (light) and B (blue to green), since the component A proved to be non-informative to the segmentation. The segmentation method in the LB space can be enhanced by adding a third dimension that is equal to the norm of the pixel-wise subtraction between the current image and the first image of the experiment (D). Then, the proposed LBD space could be segmented using any common algorithm, such as the k-means algorithm. Nonetheless, for the present study, we found that similar segmentation results could be found using a simpler approach: using the LB space, and reducing the image to an intensity array by computing L-B and normalizing the result to the interval [0,1]. Advantages were that the run-time of the algorithm was faster, and that there was no need to broadcast the initial image of the test (to compute D) to every processor when using parallel computing. After the segmentation stage, the intensity

image was binarized using the adaptive method for dark foregrounds proposed by Bradley and collaborators [144].

Nonetheless, fading thin veins could not be captured by the binarization process. This limitation did not affect considerably the overall area calculation of the slime mold cell, but it impaired the calculation of connectivity parameters. Therefore, we coupled the binarization process with an edge detection subroutine that blurred the image by using a Gaussian filter ($\sigma = 2$) and then performed ridge detection on the resulting array using a watershed procedure [145]. This subroutine yields an image that better identified thin veins but not large slime mold regions. The union of the segmentation and ridge-detection arrays gave a binary image that combined the benefits of both methods, and was then used in the subsequent steps. A sample segmented image is shown in Figure 3.1.

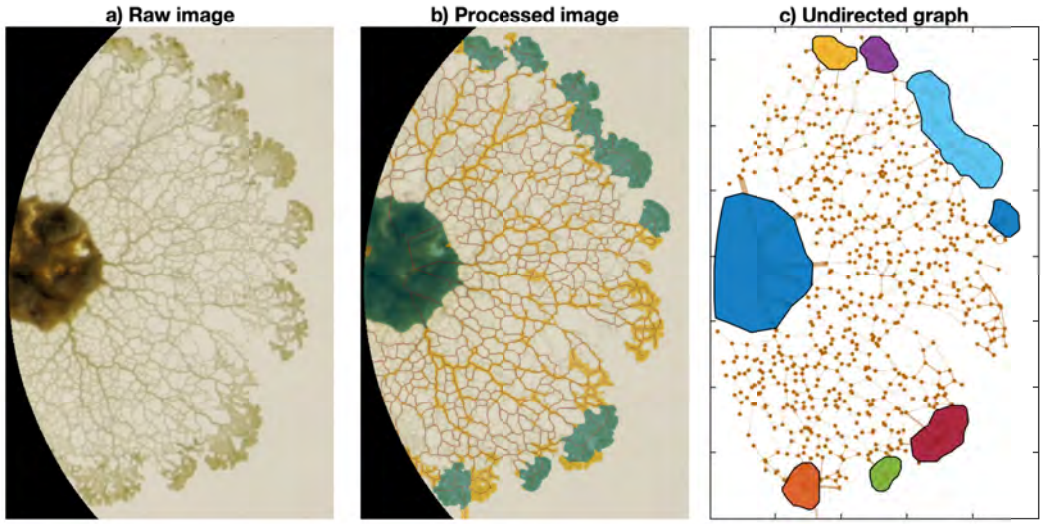

Figure 3.1: Segmentation Process. a) Raw Image, masked at the boundary of the petri dish and enhanced used a Laplacian filter. b) Segmented image overlay on original image, with highlighted pseudopods and cell skeleton. c) Resulting undirected graph: dots and segments correspond to nodes and edges respectively. Approximate location of clusters shown as colored polygons.

3.3.5 Morphological and graph network indexes - Before fusion

In the present study, we analyzed the influence of different substrates on the evolution of the slime mold network, initially focusing on each individual slime mold and later studying the

network dynamics at the location where the slime molds fuse. In both studies, morphological an topological indexes were computed in order to quantify the observations. Moreover, in order to reduce the run-time of the algorithm, we initialized the fusion time manually, by visual inspection. And then, the algorithm refined the fusion time by finding the earliest image in which the two slime molds became connected to form a single object.

Morphological indexes

The pre-fusion analysis was performed by comparing networks extracted after the slime mold cell had covered a certain area (Covered Area CA) relative to the initial slime mold area (circle of 13mm in diameter). In total, we compared networks for seven normalized areas, from 1 to 4 times the initial area of slime mold, at 0.5 intervals. The time at which the slime molds reached a given normalized area was found by using linear interpolation between known time-area pairs. Initially, the only known areas were the initial area ($t = 0$) and the area at the fusion time, which was identified manually. In the fusion analysis, we compared networks from the time of fusion, up to three hours afterwards.

Besides the slime mold cell area (CA), measured from the binary image, the so-called print area (PA) was calculated as the total area enclosed by the cell contour, e.g. including the empty regions inside it. The number of empty regions inside the slime mold cell was computed as well, together with the total area covered by these regions (PA - CA), which corresponds to the empty space inside the slime mold cell. The ratio of empty space to slime mold area (CA) has a value of 0 for a slime mold cell with no empty regions or "holes", and increases with the empty space area. The mean size of the empty regions was calculated as the ratio between the total empty area and the number of such regions.

The total length of the network was found from the topological skeleton of the segmented image, also known as medial axis, following Matlab's implementation of the algorithm proposed by Lee and collaborators [146], as highlighted in red in Figure 3.1, panel b. The average vein width was then approximated as the ratio between the slime mold

area and the total length of the network. In addition, the algorithm proposed by Maurer [147] allowed us to calculate the Euclidean distance transform along the skeleton, which is the distance from the skeleton (medial axis) to the closest edge of the slime mold cell, and therefore corresponds to half the vein width. The pseudopodia of the slime mold cell were identified by sequentially eroding and dilating the image with a structuring element of $1mm$ in radius. The remaining regions after such an operation were then labeled as pseudopods (shown in green in Figure 3.1. The proportion of pseudopods in the networks is defined as the ratio between the area of the pseudopods regions and the total slime mold area.

Undirected graph indexes

From the binary image, we obtained the skeleton and the distance transform along that skeleton. From those, we constructed undirected graphs to represent the slime mold networks. In these graphs, the nodes correspond to the branch points and the end points of the skeleton, while the segments between them correspond to the edges. Each edge is associated to its parent nodes. The edge width is equal to the average value of the vein widths along it, which was calculated as two times the distance transform value. We also calculated the length of each edge, the Euclidean distance between the parent nodes, and the drag of the edge. The drag can be thought of as the resistance to flow: it was calculated as the ratio between the edge length and the fourth power of the edge width.

We stored the coordinates of each node and we assigned a type to each node, depending on whether the node fell inside a pseudopod, the initial cell, or the veins. After constructing the graph, a subroutine calculated the connectivity of each node, and simplified the graph in order to avoid self-loops and repeated edges between nodes. A sample undirected graph constructed from a raw image is shown in Figure 3.1, panel c. Four different categories of graph indexes were computed: connectivity, cost, transport efficiency and resiliency/hierarchy, as explained below. The indexes are illustrated in the simple sample

graph shown in Figure 3.2.

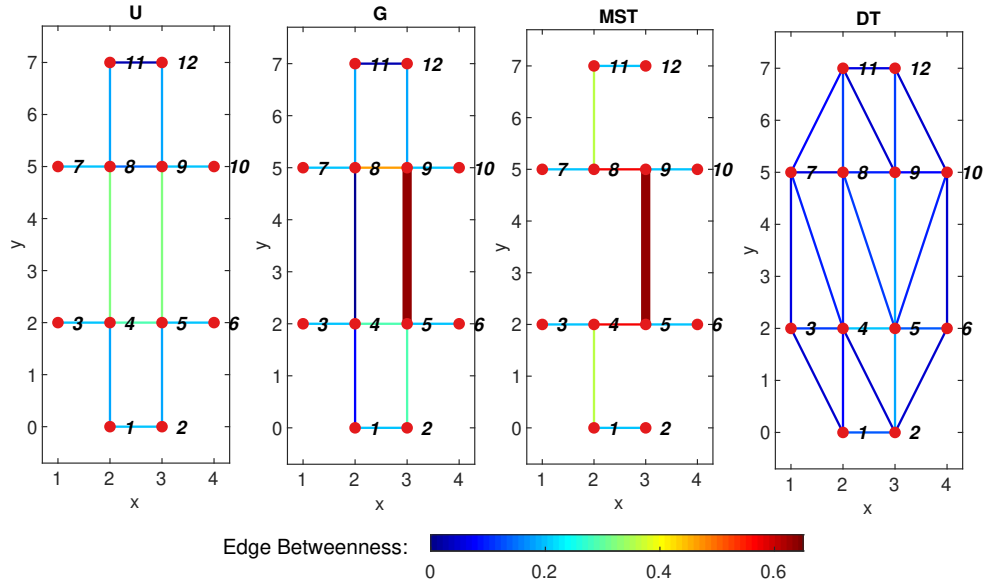

Figure 3.2: Sample graph. U: sample undirected graph connecting 12 nodes, all the edges in the graph have a uniform edge width. G: sample graph with the same topology as graph U, but an increased edge width between nodes 5 and 9. MST: Minimum spanning tree for U and G, some edges disappear from the graphs to form a simply connected network with minimum length. DT: Delaunay triangulation of graphs U and G, the triangulation is a maximum connected graph between the nodes, irrespective of the edge or node weights of the graph. The edge color is proportional to the edge betwenness in the graph.

Connectivity

The node degree corresponds to the number of edges connected to a given node. It is a common metric used to quantify the ”connectedness” of a network. Nevertheless, biological networks tend to form degree 3 networks with little variations, and therefore the mean node degree did not provide much information on network connectivity. The alpha index [148] is a measure of the density of cycles in the graph (also named loops), ranging from 0 for simply connected networks (no loops) to 1 for maximum planar networks, which are networks in which no extra edges can be added without crossing existing ones (e.g. Delaunay triangulations). The alpha index was computed but, similar to the connectedness index, results were non-informative.

Network Cost

We quantified the network cost in terms of the weighted wiring cost under different scenarios. The formal definition of the wiring cost is the sum of the upper (or lower) triangular weighted adjacency matrix of the graph. If the edge weight corresponds to the length of the edges, the wiring cost corresponds to the total length of the network. Instead, if the weight metric is the Euclidean distance between connected nodes, the wiring cost corresponds to the total length of a network with the same topology, but with strictly straight segments between nodes. Then, we defined the tortuosity of the network as the ratio between the total network length (wiring cost by length) and the length of the equivalent graph with straight edges (wiring cost by Euclidean distance).

Furthermore, we normalized the total network length by comparing it to its upper and lower bounds. The minimum spanning tree (MST) is the subset of edges of the original graph that connects all the nodes with minimum total edge weight, and therefore is a lower bound for network cost. For instance, the MST in terms of length is the third graph in Figure 3.2. In this example, the MST of U in terms of length or drag is not unique, given that the graph has double symmetry and that there is more than one solution that gives the minimum network length/drag. For instance, choosing either edge from 4-8 or 5-9 will yield the same (minimum) network length/drag. On the contrary, the MST in terms of drag for graph G strictly selects edge 5-9 over edge 4-8.

The upper bound for network cost is the Delaunay triangulation (DT), which is the maximally connected graph for the set of nodes in the graph, irrespective of weight measures. A maximally connected graph is such that there is no way to add an extra edge to the network without crossing any other existing edge. Given that the DT of a graph depends only on the geometric distribution of its nodes, not the weight of the edges, the triangulation (DT) shown in Figure 3.2 is the same for graphs U and G. Once we found the comparison graphs for each network, we mapped the network cost to the interval [0, 1] where zero and one correspond to the lower and upper bounds, respectively.

Transport efficiency

Transport efficiency was quantified based on the travel distance (shortest path) and the drag between source and sinks. Similar to the network cost, we compared the transport efficiency between pairs of nodes in the network against a lower bound. We computed the shortest path between every pair of nodes along the network (using Matlab's implementation of Dijkstra's algorithm) in terms of edge length, and its lower bound, the Euclidean distance between node coordinates. Then, we calculated the length efficiency index (LE) defined [149] as follows:

$$LE = \frac{\sum_{ij} l_{ij}^{-1}}{\sum_{ij} e_{ij}^{-1}} \tag{3.1}$$

Where l_{ij} and e_{ij} are the path length and the Euclidean distance between nodes i and j, respectively.

Referring to the sample graphs in Figure 3.2, one can see the differences in transport efficiency, for instance between nodes 2 and 7 of the graphs. It becomes evident that the DT offers the best efficiency, since the path from node 2 to 7 passes only through nodes 2, 4 and 7, with a total path length that is similar to a straight line between nodes 2 and 7. Conversely, the same path along graph U passes through nodes 2, 1, 4, 8 and 7 (it is not unique), with a path length that is clearly longer than a straight line, and therefore, with lower efficiency. Moreover, the same path along the MST passes through nodes 2, 1, 4, 5, 9 and 8, yielding an even longer path length and lower efficiency, showing that network cost and path efficiency are oftentimes inversely correlated.

In addition to the general efficiency index (LE), we also studied the distribution of path efficiency, by computing e_{ij}/l_{ij} for each pair of nodes i, j and by calculating the mean value and the coefficient of variation (COV, standard deviation divided by the mean) of the set of individual path efficiencies.

Similarly, we computed the pairwise shortest paths in the graph, this time in terms of

minimum drag between nodes d_{ij}. In this case, the lower bound is a path with minimum possible drag between nodes, which corresponds to the shortest path (straight line) through an infinitely wide vein, which equals zero. For comparison, our lower bound c_{ij} was the drag between nodes i, j through a vein with a diameter of 0.5 mm and shortest possible length (euclidean distance). Setting an arbitrary vein width (in this case 0.5 mm) can shift and stretch the values of drag efficiency, but it does not modify its distribution, and therefore, it does not alter the relative values among graphs. We thus computed the mean and COV of the drag efficiencies d_{ij}/c_{ij} for every graph, and we compared the values relative to one another instead of discussing the significance of their magnitudes.

Edge hierarchy and network resiliency

In order to quantify the relative importance of edges in the graph, we calculated the edge betweenness centrality, defined as the percentage of all the shortest paths between nodes that pass through each edge. Similar to the shortest path efficiency, we calculated the edge betweenness in terms of length and drag. An example of the edge betwenness distribution is shown in Figure 3.2 in the form of edge colors. From the DT, it is evident that each edge is used by approximately 8% of the shortest paths between nodes, with a very homogeneous distribution. On the contrary, more centralized networks such as the MST, show the dependency of the network on some edges, such as edge 5-6 which is used in about 65% of the shortest paths, since it is the only node connecting the upper and lower regions of the graph. A comparison between the drag betweenness of graphs U and G shows the influence of the edge weight (drag in this case), even though the topology of both networks is the same, the low drag of edge 5-9 in graph G skews the betweenness distribution, showing that this edge becomes significantly more important in the graph in comparison to edge 4-8. Apart from the difference in drag between those edges (graph U), the load is distributed and becomes symmetric at about 32%. The frequency distribution of betweenness for the graphs followed an exponential distribution for both metrics (length and drag); it was parametrized

by the parameter μ, equivalent to the mean of the edge betweenness values.

In addition, we studied the network resiliency in terms of the fault tolerance (FT) of the graphs [150, 25, 149]. The FT corresponds to the percentage of edges that can be removed from the graph while still connecting a given percentage of the nodes, usually 50%. The selection of the edges to remove was random and so the FT was calculated as the average of 30 independent realizations of the procedure. In order to provide bounds of comparison, we calculated the FT for the MST (lower bound) and the DT (upper bound) of the graph. We mapped the FT of the actual network to the interval [0,1] where the bounds correspond to the MST and the DT, respectively.

Moreover, in order to study the effect of edge hierarchy in the resiliency of the network, we proposed a modified procedure to calculate the FT, this time by systematically removing the k edges with the highest/lowest drag betweenness with $1\% \leq k \leq 100\%$ of the number of edges in the graph. Removing the most 'important' edges of the graph (highest betweenness) is considered a worst-case scenario, and starting from the least important edge, a best-case scenario.

3.3.6 Cell fusion indexes

We tracked the fusion region from fusion time (FT) up until 3 hours after fusion, by analyzing images every 15 minutes. In order to find the fusion region, we compared the print of both slime mold cells right before fusion (BF), and the print of the fused slime mold five minutes after fusion (AF). The difference between AF and BF (AF-BF) gave an image of the newly grown regions (r_i), from which, one (or more) vein connected the two slime mold cells. We calculated the region obtained by adding each r_i to BF until the regions became connected. Then, at FT, we extracted the region of slime mold that was within 5 mm from the connecting region. In order to capture possible growth of slime mold adjacent to this region, we added an offset of 2 mm around its contour. The resulting area was defined as the fusion region (FR).

Then, at each time step, we measured the slime mold area, print area, ratio of empty and slime mold areas, average size of the enclosed empty regions, network length and mean and maximum vein widths inside the fusion region (FR). We observed large variability among the index values, attributed to the wide differences among the shape and extent of the fusion regions themselves. To alleviate this issue and better observe the evolution after fusion, most of these indexes were normalized by their initial value at fusion. Indexes specific to the fusion region were defined based on graph analysis, and are explained in the following section.

Fusion region graph indexes

We analyzed the evolution of the graph within the fusion region in terms of number of nodes and edges, mean and maximum edge thickness within FR. In addition, we studied the connectivity between the two initial slime mold cells and its evolution over time, which was highly variable among environments, as shown in Figure 3.3. To do this, we calculated the shortest path between the two initial locations of the slime mold cells in terms of length and drag and in terms of the number of veins connecting the initial slime mold cells. The number of connections between the initial slime mold cells was calculated from the maximum flow of the graph, using Matlab's implementation of Boykov-Kolmogorov's algorithm. In general terms, the maximum flow of the graph between two nodes is the critical edge capacity that limits the flow between such nodes. Then, in the case of unweighted graphs (i.e. every edge has a weight of 1), the maximum flow corresponds to the maximum number of distinct paths between the nodes, i.e. the number of connecting veins.

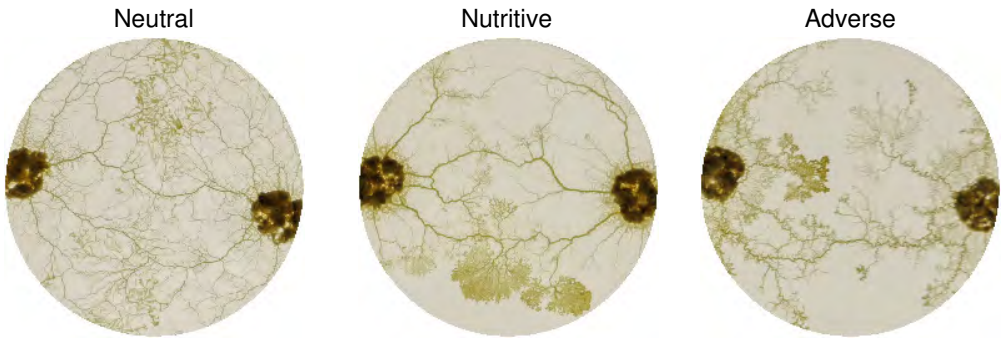

Figure 3.3: Fused slime mold cells. Raw images taken three hours after initial fusion of the slime mold cells in different substrates. Images illustrate the influence of the substrate in the number and thickness of connecting veins between sides of the network.

Then, in order to count the maximum number of veins connecting initial slime mold cells, we used an initially unweighted graph, we picked one node inside each initial cell, and increased the weight of their surrounding edges to a large value. This to make sure that the connections bottleneck occurred in the connection between both sides of the network. Specifically, we increased the weight of the edges within 25 mm of each source node to a value of 100. Then, the maximum flow between the nodes corresponded to the number of connecting veins between the initial slime mold cells. Furthermore, following the same procedure with a graph weighted by its edge width, we calculated maximum flow as the total bandwidth between both sides of the graph. We present the results both in terms of number of connecting veins and in terms of the average width of the edges, which was calculated as the total bandwidth divided by the number of connecting veins. In a similar way, the shortest path between the initial slime mold cells was tracked over time, both in terms of minimum path length between initial slime mold cells and in terms of minimum drag path.

3.3.7 Statistical analyses

To assess the difference in the various parameters measured between the three treatments, we used linear mixed models (function lmer, Package lme4) or generalized linear mixed

models (function glmer, Package lme4 [151]) in R (RStudio Version 1.2.1335). The models were fitted by specifying the fixed effects: treatment (categorical predictor), the normalized area or the time after fusion (continuous predictors) and the random effects: the plasmodium and the replicate. The dependent variables that did not fit linear model requirements were transformed using the "bestNormalize" function ("bestNormalize" package [152]). The outcomes of all the models are presented in the supplementary information published in [28].

3.4 Results

The following section is divided into two parts. First, we study the influence of the different environments (neutral, adverse and nutritive) on the network characteristics of the slime mold cells, growing freely before fusion. Secondly, we analyze the network dynamics after fusion in the different environments.

3.4.1 Influence of substrate on slime mold networks

slime mold morphology Analysis

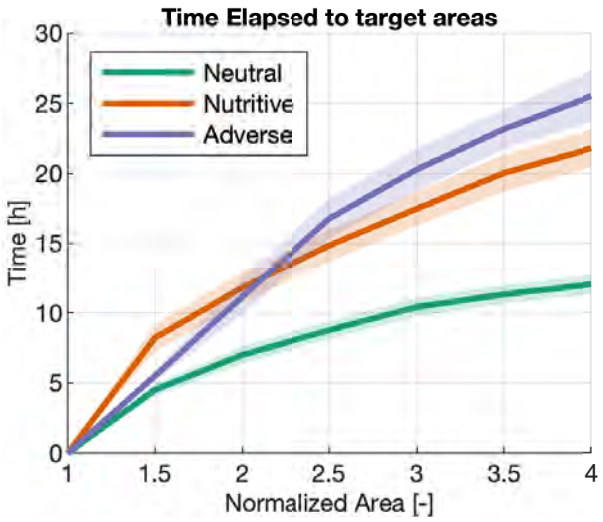

Figure 3.4: Time elapsed to target areas. Slime mold area normalized by the initial cell size (horizontal axis) vs. time elapsed to reach such area (vertical axis). Solid lines correspond to the mean value among replicates for the same substrate and shaded regions correspond to the confidence interval.

The first step of the analysis focuses on the growth rate of the slime mold cells in different environments, as shown in Figure 3.4. Results show that slime molds exploring a neutral environment expand to cover four times their initial area after 10 hours whereas slime molds exploring a nutritive or an adverse environment reach the same area only after 21 and 26 hours in average, respectively (Figure 3.4). This result confirms that the expansion rate depends on the substrate characteristics, as observed in our previous study [27].

Note that the total area enclosed by the whole slime mold cells is not always the same. Therefore, we quantified the ratio between the empty area (not covered by slime mold) and the slime mold cell area, as shown in Figure 3.5.

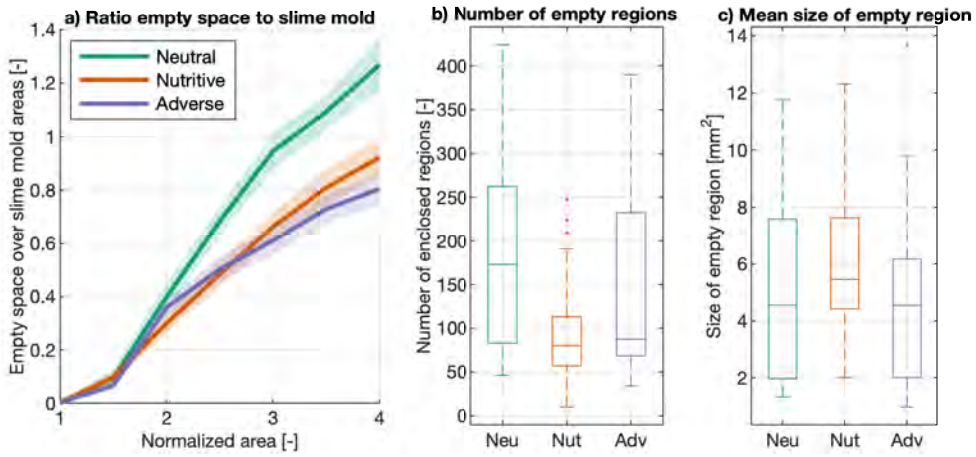

Figure 3.5: Empty regions inside slime mold cell. a) Ratio between empty space inside slime mold cell and slime mold area as a function of normalized slime mold cell area (An). The solid lines and shaded areas correspond to the means and confidence intervals among replicates, respectively. b) Boxplot graph of the number of enclosed empty regions inside the slime mold cell at An=4. c) Boxplot graph of the mean size of the empty region at An=4, computed as the total empty area divided by the number of enclosed regions. Box plots show median (horizontal line), interquartile range (box), distance from upper and lower quartiles times 1.5 interquartile range (whiskers), and outliers ($> 1.5x$ upper or lower quartile)

While expanding on a substrate, slime molds build networks that enclose empty regions and appear as mesh net structures. The total area covered by such empty regions is larger when a slime mold is exploring a neutral substrate than the other two substrates (Figure 3.5a). Since we are comparing the geometry and network of slime mold cells that cover the same area (CA), it also means that the total area explored by slime molds (print area PA) in the neutral environment is larger than on the other two substrates. The number of empty regions is the lowest on a nutritive substrate and the highest on the neutral substrate (Figure 3.5b) while the area of the empty regions is the smallest on the adverse substrate (Figure 3.5c). Thus, on a neutral substrate, the slime molds build sparse networks enclosing numerous empty spaces, while on a nutritive substrate, they construct networks presenting few but large empty regions, and on an adverse substrate, they establish a tight and compact network. We also notice that the proportion of pseudopods is the highest on a nutritive substrate.

slime mold Network Topology Analysis

The topology analysis of the slime mold networks first shows that on all substrates, the number of edges and the number of nodes are highly correlated regardless of the network size. The mean node degree is in average 2.65. Figure 3.6 shows that there is a linear relationship between the number of nodes and edges in the networks, which holds for every environment and cell size studied. The fact that the number of edges is consistently 35% more than the number of nodes explains that the connectivity of the networks is an inherent characteristic of the slime mold networks.

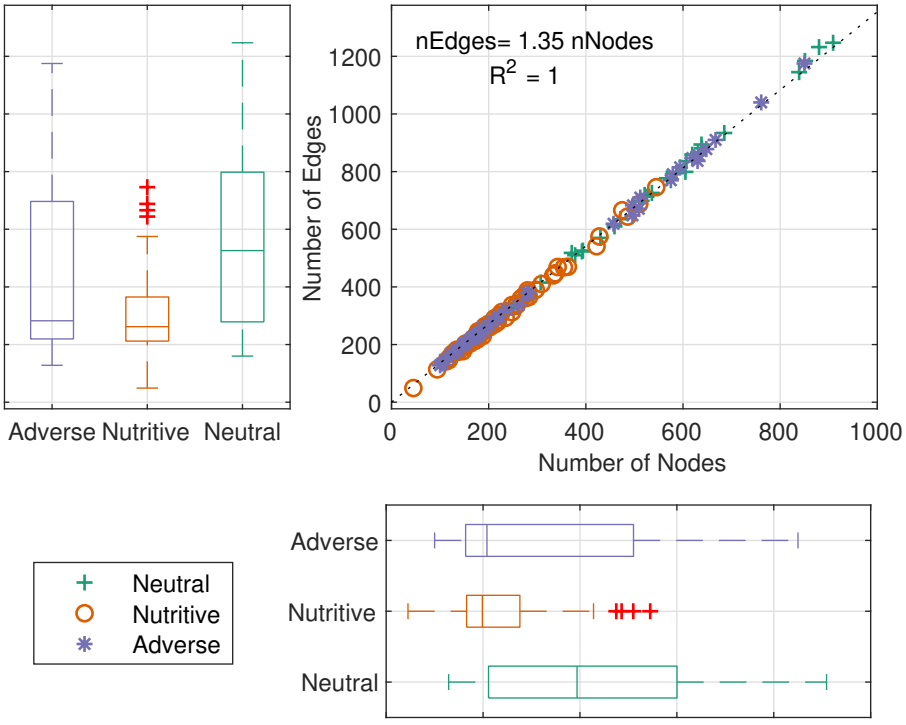

Figure 3.6: Number of nodes and edges in the network. Boxplots showing the variability of the number of Edges and Nodes in the networks by substrate. Scatter plot showing the linear relationship between the number of edges and nodes. Box plots show median (horizontal line), interquartile range (box), distance from upper and lower quartiles times 1.5 interquartile range (whiskers), and outliers ($> 1.5x$ upper or lower quartile)

Regarding the total length of the networks, slime molds migrating on a neutral substrate build longer networks (Figure 3.7a) with narrower veins (Figure 3.7b, right hand-side and

Figure 3.7c), more nodes (Figure 3.6) and more edges (Figure 3.6) than on the other two substrates.

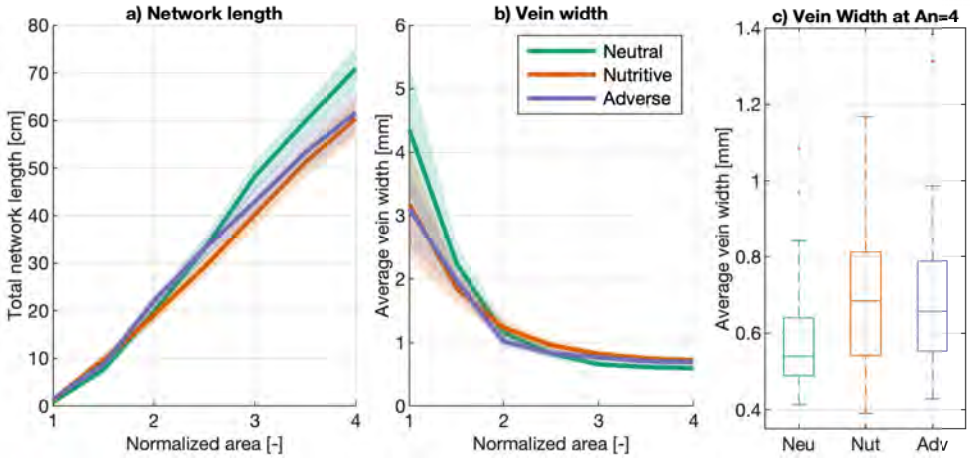

Figure 3.7: Total Network Length and average vein width. a) total network length as a function of normalized slime mold cell area (An). b) Average vein width of the network as a function of normalized slime mold cell area (An), computed as the ratio between the slime mold area and the total network length. c) Boxplot graph of the average vein width at An=4. The solid lines and shaded areas correspond to the means and confidence intervals among replicates, respectively. Box plots show median (horizontal line), interquartile range (box), distance from upper and lower quartiles times 1.5 interquartile range (whiskers), and outliers ($> 1.5x$ upper or lower quartile).

We observe that when slime molds reach three times their initial area (An=1), the vein width becomes constant. We also note a general decrease of vein width over the initial growth from An=1 to An=3, which was expected, since the biomass remains constant in these environments [27]: pseudopods are gradually transformed into veins. The main interpretation is that on a neutral substrate, the print area (PA) of the slime molds is larger because the slime molds are exploring: there are more empty spaces, and there are more connections between nodes and between edges. Since we are comparing slime molds that have the same cell area (CA) (and density), veins have to be narrower than in the nutritive and adverse environments, by biomass conservation (Figure 3.7).

The analysis of vein width and vein length distribution indicates that veins are the narrowest and the shortest when slime molds explore a neutral environment (Figure 3.8), following a log-logistic distribution with abundant short and thin veins and few long and thick

veins.

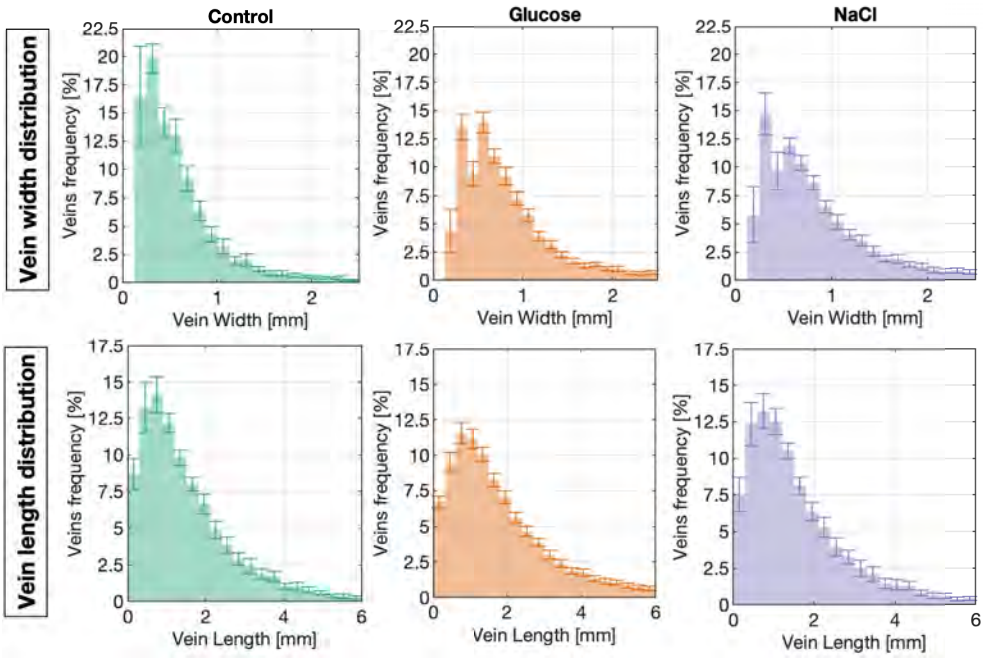

Figure 3.8: Distribution of veins width (top row) and length (bottom row). Bars in the frequency histograms show the mean percentage of network veins within a given range, while vertical error bars show the confidence interval.

The analysis of the normalized length of the networks shows that in general, slime mold networks have low tortuosity (e.g. edge lengths are close to the straight-line distances between nodes), with a slight difference in tortuosity and normalized network length between slime molds exploring a neutral substrate and slime molds exploring a nutritive one, the latter being the most tortuous (Figure 3.9a), the former being the longest (Figure 3.9b). Interestingly, we observe that the total length of the slime mold network is closer to that of a minimum spanning tree (which corresponds to a normalized length of zero) than to the total length of a fully connected network of the same topology (Delaunay triangulation, normalized length of 1), which at the same time, suggests that even though slime molds form networks with loops, those networks are far from being fully connected networks, keeping their total length closer to the MST.

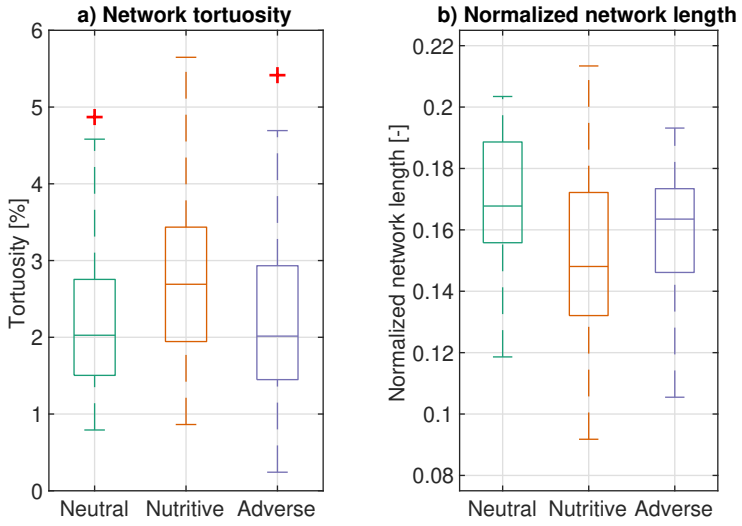

Figure 3.9: a) Network tortuosity, computed by comparing the length of the network to the length of a network of equivalent topology but strictly straight segments between nodes. b) Normalized network length, mapped to the interval [0 1], where the minimum spanning tree (MST) and Delaunay triangulation DT are the lower and upper bounds respectively. Box plots show median (horizontal line), interquartile range (box), distance from upper and lower quartiles times 1.5 interquartile range (whiskers), and outliers ($> 1.5x$ upper or lower quartile)

The efficiency of the network measures how effective the network is to connect all its regions with one another. We calculate such efficiency in terms of path length (length efficiency, LE) and in terms of path drag (drag efficiency, DE). Drag can be understood as the resistance to flow between two points, and therefore a higher drag efficiency means higher ease of flow. The length efficiency is the lowest for the slime molds exploring a nutritive substrate (Figure 3.10a) whereas the drag efficiency is the lowest for the slime molds exploring a neutral substrate (Figure 3.10b).

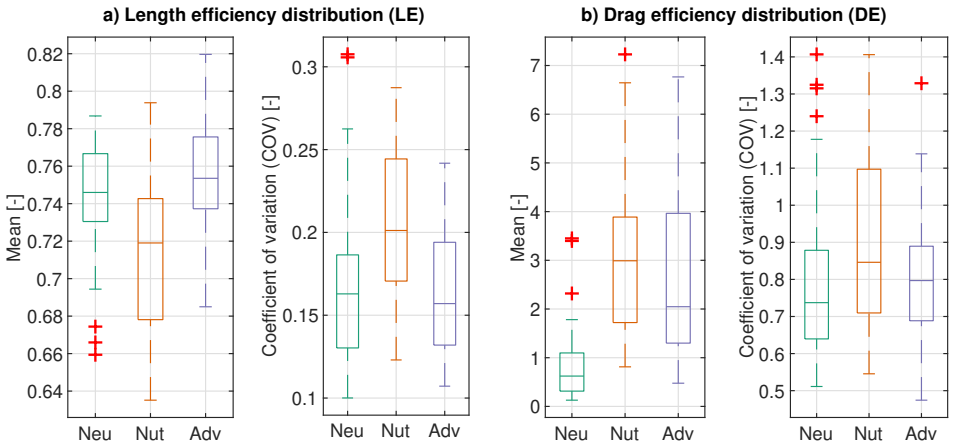

Figure 3.10: Transport efficiency. a) Length efficiency distribution: ratio between the distance along the network and the Euclidean distance between every pair of nodes in the network. b) Drag efficiency distribution: ratio between the drag along the network and along a straight path with a 0.5mm width (arbitrary, constant value). The coefficient of variation (COV) corresponds to the standard deviation divided by the mean of the values. Box plots show median (horizontal line), interquartile range (box), distance from upper and lower quartiles times 1.5 interquartile range (whiskers), and outliers ($> 1.5x$ upper or lower quartile)

Another important metric of transportation networks is centrality, which is a measure of importance of the nodes or edges in the network, in terms of how often they are used during transport between locations. A homogeneous network has an even distribution of node/edge importance while a centralized network depends heavily on some important nodes/edges to route the network traffic. In the following, we measure edge importance in terms of betweenness centrality relative to both length and drag, which we calculate as the percentage of shortest paths (by length or drag) that pass through each node. The distribution of node importance in the networks (represented by mean edge betweenness), shows that slime molds in neutral environments are more homogeneous than those exploring nutritive and adverse environments. Slime molds in a nutritive environment are the most centralized among treatments (Figure 3.11 a-b).

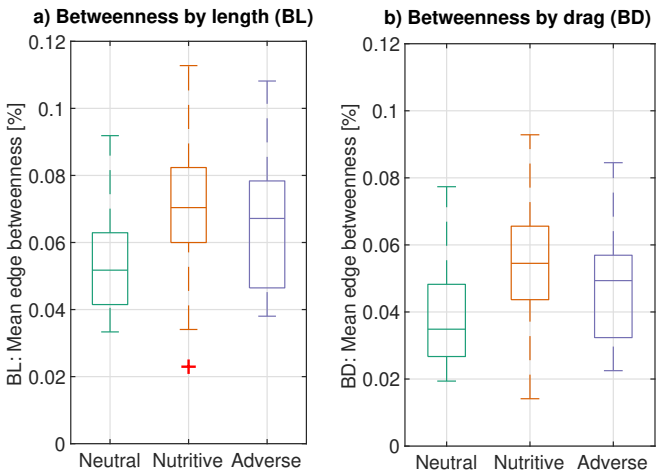

Figure 3.11: Edge betweenness: measure of edge importance, calculated as the percentage of shortest paths between nodes that pass through each edge. a) Mean betweenness considering that the graph is weighted by the edge length (BL). b) Mean betweenness considering that the graph is weighted by the edge drag (BD). Box plots show median (horizontal line), interquartile range (box), distance from upper and lower quartiles times 1.5 interquartile range (whiskers), and outliers ($> 1.5x$ upper or lower quartile)

Network resiliency, understood as the capacity of a network to perform adequately even with damage, is measured in terms of normalized fault tolerance. Fault tolerance is calculated as the percentage of randomly chosen edges that can be removed from the network while still connecting 25%, 50% or 75% of the nodes in the network. Fault tolerance is normalized to the interval [0, 1], where the bounds correspond to a simply connected (MST) and a maximally connected (DT) networks respectively. Results from Figure 3.12 show that the networks in the neutral and adverse environments are similar in terms of resiliency, while the networks in the nutritive substrate are the least resilient (Figure 3.12). The fact that networks in a nutritive environment are the most centralized and the least resilient suggest that they depend heavily on certain edges which sustain the connectivity of the network. On the contrary, the exploratory behaviour of slime molds in neutral or adverse conditions yields networks that are less centralized and more resilient, e.g. more capable to overcome accidental disconnections.

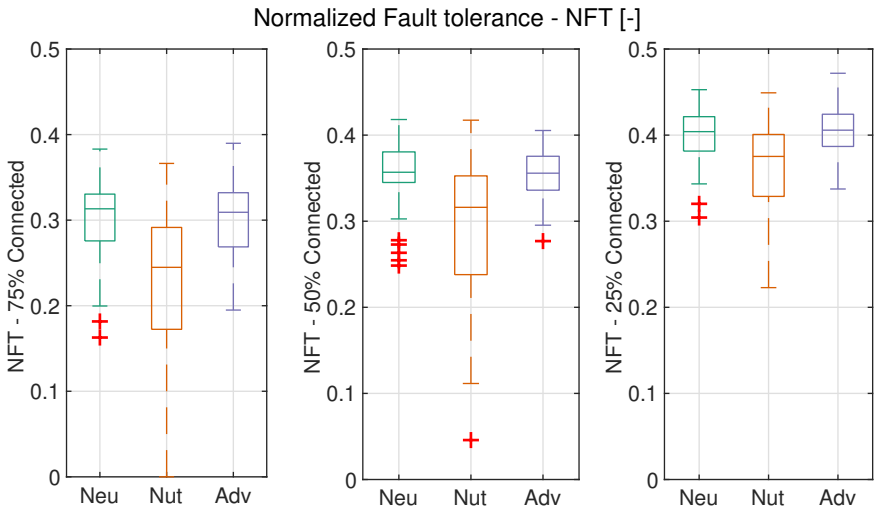

Figure 3.12: Normalized fault tolerances. The random fault tolerance (FT) is calculated as the percentage of edges that can be removed from a graph while still being able to connect 25% (left), 50% (middle) or 75% (right) of the graph edges. The FT is normalized to the interval [0, 1] where the lower and upper bounds correspond to the minimum spanning tree (MST) and Delaunay triangulation (DT), respectively. The edges to remove are chosen randomly and therefore the average of 30 replicates is reported for each graph and its bounds. Box plots show median (horizontal line), interquartile range (box), distance from upper and lower quartiles times 1.5 interquartile range (whiskers), and outliers ($> 1.5x$ upper or lower quartile)

Moreover, the relationships between network cost (normalized network length) and resiliency (normalized fault tolerance) and the edges importance distribution (mean drag betweenness) show that the normalized fault tolerance is positively correlated with the normalized network length (Figure 3.13a) whereas the mean drag betweenness is negatively correlated with the normalized network length (Figure 3.13b), Which shows that networks with higher network length have a more balanced importance distribution and are also more resilient.

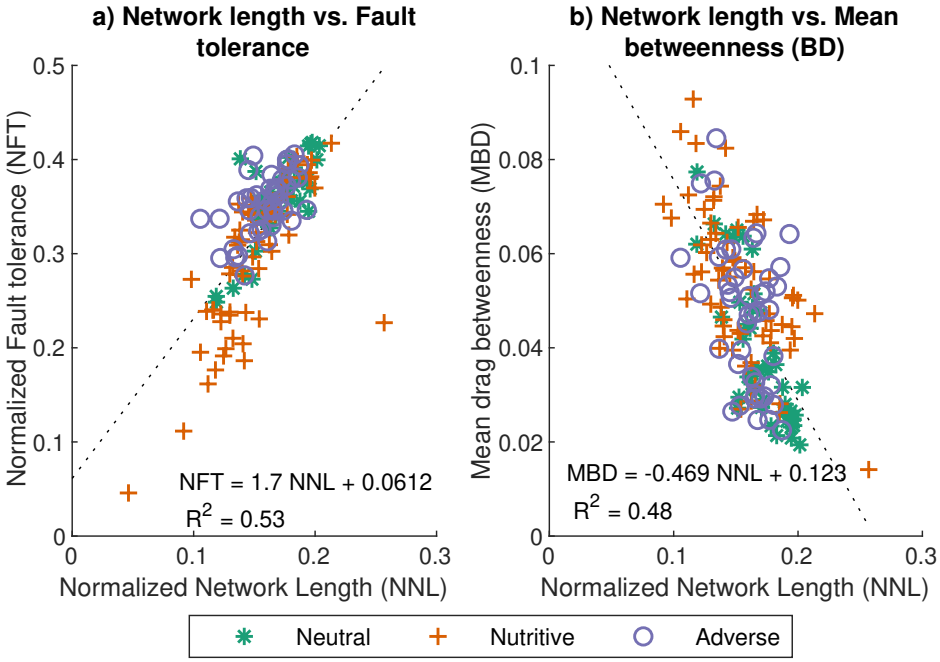

Figure 3.13: a) Relationship between normalized network length (NNL) and random fault tolerance for 50% connected nodes (FT50). b) Relationship between normalized network length (NNL) and mean drag betweenness (BD).

Given that the normalized fault tolerance (NFT) is based on random removal of edges, it does not take into account the importance of the edges in the network (also called edge hierarchy). For this reason, it is interesting to calculate the decay of the number of nodes connected as a function of removed edges, where this time, the edges are sorted by their importance (drag betweenness) and then removed from the least (most) to the most (least) important edges, which correspond to the best (worse) case scenarios respectively for network resiliency. Results are shown in Figure 3.14 together with the mean values of random fault tolerance shown in Figure 3.12 (before normalization).

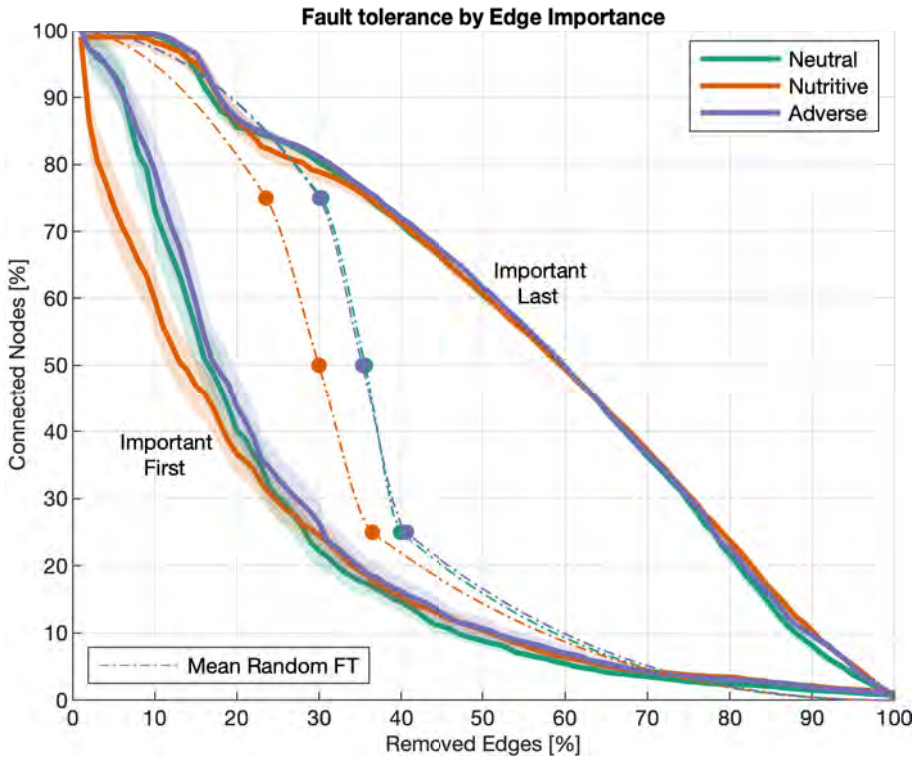

Figure 3.14: Fault tolerance Decay. Fault tolerance by edge importance. Decrease of the percentage of connected nodes in the graph (vertical axis) as a function of edge removal (horizontal axis). Dashed lines correspond to the mean random tolerance calculated for 25 %, 50 % and 75 % of the nodes, connected by a spline. Solid lines and shaded regions correspond to the mean and confidence intervals (CI) for the worst and best case scenarios, resulting from removing the most important edges first and last, respectively.

Results show that edge hierarchy significantly influences the fault tolerance of networks, showing a steep loss of connected nodes (around 70% disconnected edges) after the 25% of the most important edges are removed from the networks. This effect is even more considerable on the nutritive substrate, which is consistent with the fact that these networks are more centralized than those on the other substrates (see Figure 3.11). However, regardless of the substrate, the percentage of connected edges remains almost unchanged (above 95% of connected nodes) even after removing the 15% of the edges with the lowest importance, which suggests that these edges are not essential to the flow of the network, but significantly contribute to resiliency. Regarding the random fault tolerance, the networks on neutral and adverse substrates show very similar resiliency while the networks on the

nutritive substrate are more vulnerable to damage.

3.4.2 Analysis of slime mold fusion process

The second part of the analysis focuses on the fusion process between two slime mold cells growing on different substrates. We measured the time at which the slime molds fused and the area of the whole slime mold as soon as they fused. Networks on a neutral environment reach fusion faster, which is consistent with the growth rates observed in Figure 3.4. Nonetheless, the area of the slime mold cells at fusion exhibits no significant differences among treatments, showing that the differences among treatments correspond to difference in growth rates and not in slime mold cell morphology.

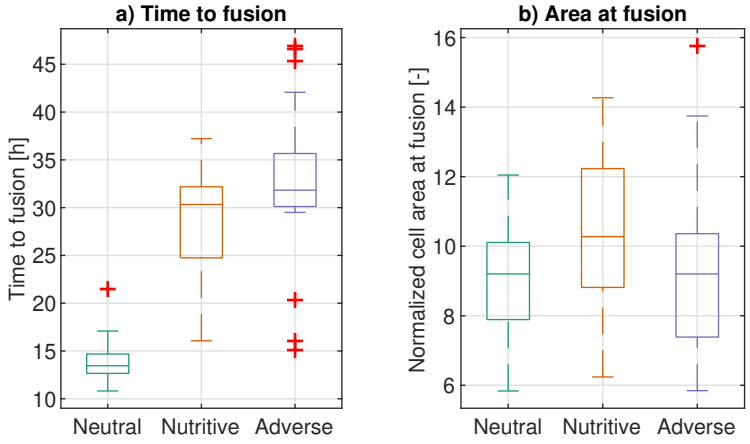

Figure 3.15: Time to fusion. a) Boxplot with distribution of time elapsed from the beginning of the experiment until fusion between slime molds occurred. b) Distribution of normalized areas (relative to initial cell area) of the slime mold cells at the time of fusion. Box plots show median (horizontal line), interquartile range (box), distance from upper and lower quartiles times 1.5 interquartile range (whiskers), and outliers ($> 1.5x$ upper or lower quartile).

In the following, we focus on the fusion region, i.e. the neighborhood where the slime mold cells fused at the first place. We tracked its characteristics for three hours from the fusion time. First, we calculated the number of edges and nodes inside the fusion region. Figure 3.16 shows the relationship between nodes and edges within the whole network, and just inside the fusion region.

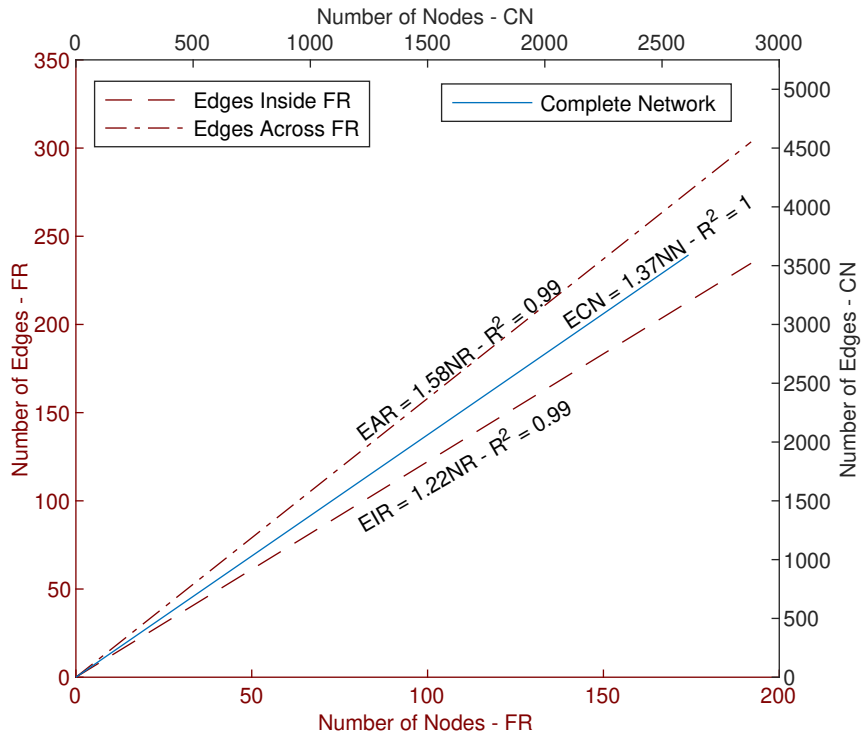

Figure 3.16: Number of nodes and edges inside the fusion region and the full network. Dashed lines (left and bottom axis) correspond to the relation between the number of nodes inside the fusion region (FR) and the number of edges inside it, e.g. both parent nodes of the edge lie inside FR, and number of edges across FR, i.e. at least one of the parent nodes of the edges lies inside FR. Solid line (right and top axis) correspond to the total number of nodes and edges in the complete, fused network.

Results show that the relationship between nodes and edges in the whole network is similar to that observed in the individual slime mold cells (Figure 3.6). Within the fusion region, this relationship depends on the way the edges inside the region are counted. But in any case, the node-to-edge relationship has a strong correlation irrespective of the substrate.

Next, we study the morphology evolution of the slime mold cell within the fusion region. Figure 3.17 shows the change of slime mold area, print area and enclosed area.

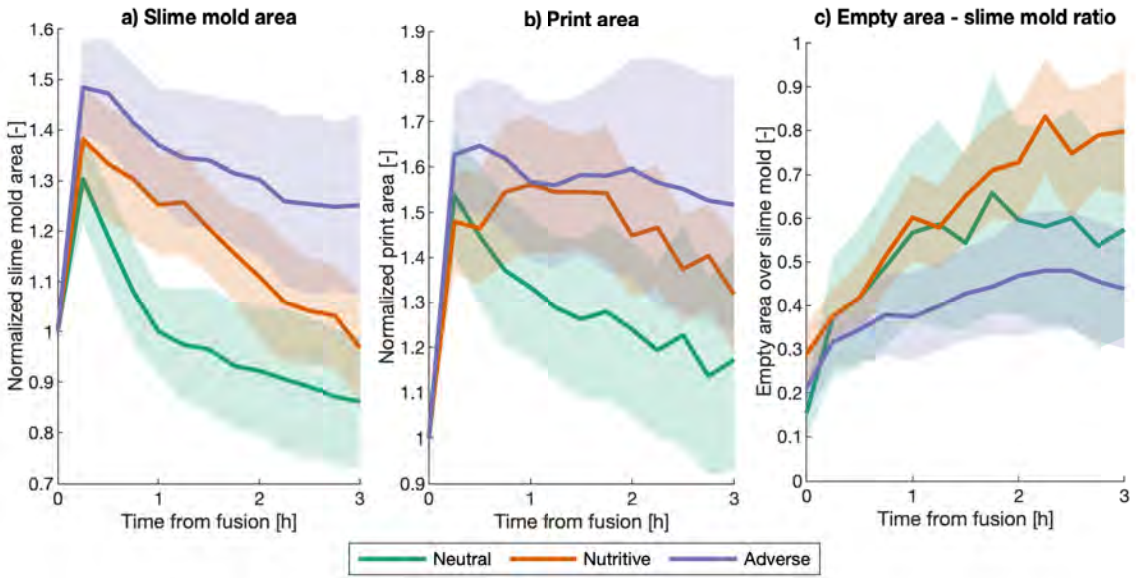

Figure 3.17: Area indexes inside the fusion region (FR). a) slime mold area normalized by the area at fusion. b) Print area of the slime mold cell (including enclosed empty space inside cell), normalized by the print area at fusion. c) Ratio between enclosed empty space and slime mold areas.

Results have been normalized with respect to their values at fusion time. In general, the slime mold and print areas increase for a certain time after fusion, reaching a maximum, and then steadily decrease. Statistic analyses show that the rate at which the slime mold area decreases after fusion on a neutral substrate is significantly higher than the rate on the adverse and nutritive substrates. Similarly, the rate at which the print area decreases in the neutral environment is higher than that in the nutritive environment. Lastly, the ratio of empty area to slime mold area increases steadily for all the treatments, suggesting that the slime mold is refining, becoming less compact in the process. Slime mold is refining at a faster rate in the nutritive environment than in the neutral environment.

Results from Figure 3.18a show the network length significantly increases after fusion, and even though it eventually reduces, it is always higher than the initial network length at fusion. The refinement rate at which the network length decreases in the neutral environment is higher than that in the nutritive environment. Refinement inside the fusion region is further observed through the evolution of the average and maximum vein widths

(Figure 3.18b-c), which reduce after fusion and eventually reach an equilibrium. Statistical analyses show that the rate at which the maximum and average vein widths decrease is faster in neutral environments compared to nutritive ones. In addition, the maximum vein widths after 3 hours from fusion are significantly lower on the neutral substrate and are significantly different in the nutritive environment compared to the other two.

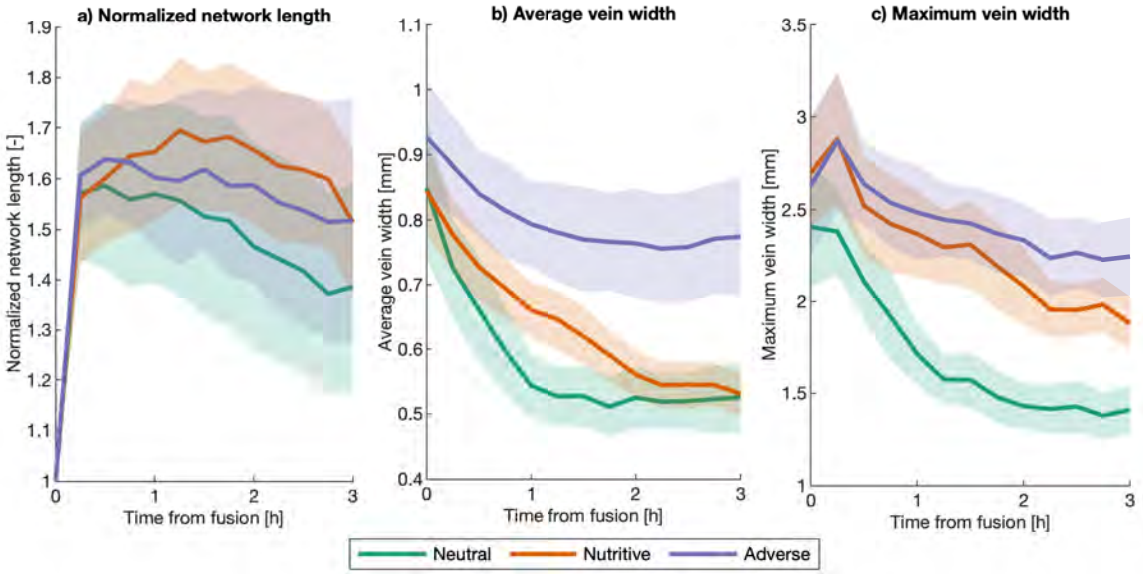

Figure 3.18: Network length and vein width inside FR a) Network length inside normalized by the network length at fusion. b) Average vein width inside FR. c) Maximum vein width inside FR. Solid lines correspond to the mean value among replicates for the same substrate and shaded regions correspond to its confidence interval.

To analyze the evolution of the connectivity between the previously separated slime mold cells, we tracked the number and average width of the veins that connect both sides of the newly fused slime mold and we also tracked the path distance and the drag between both sides. The number and width of the connecting veins are found from the maximum flow of the graph, as described in the methods section.

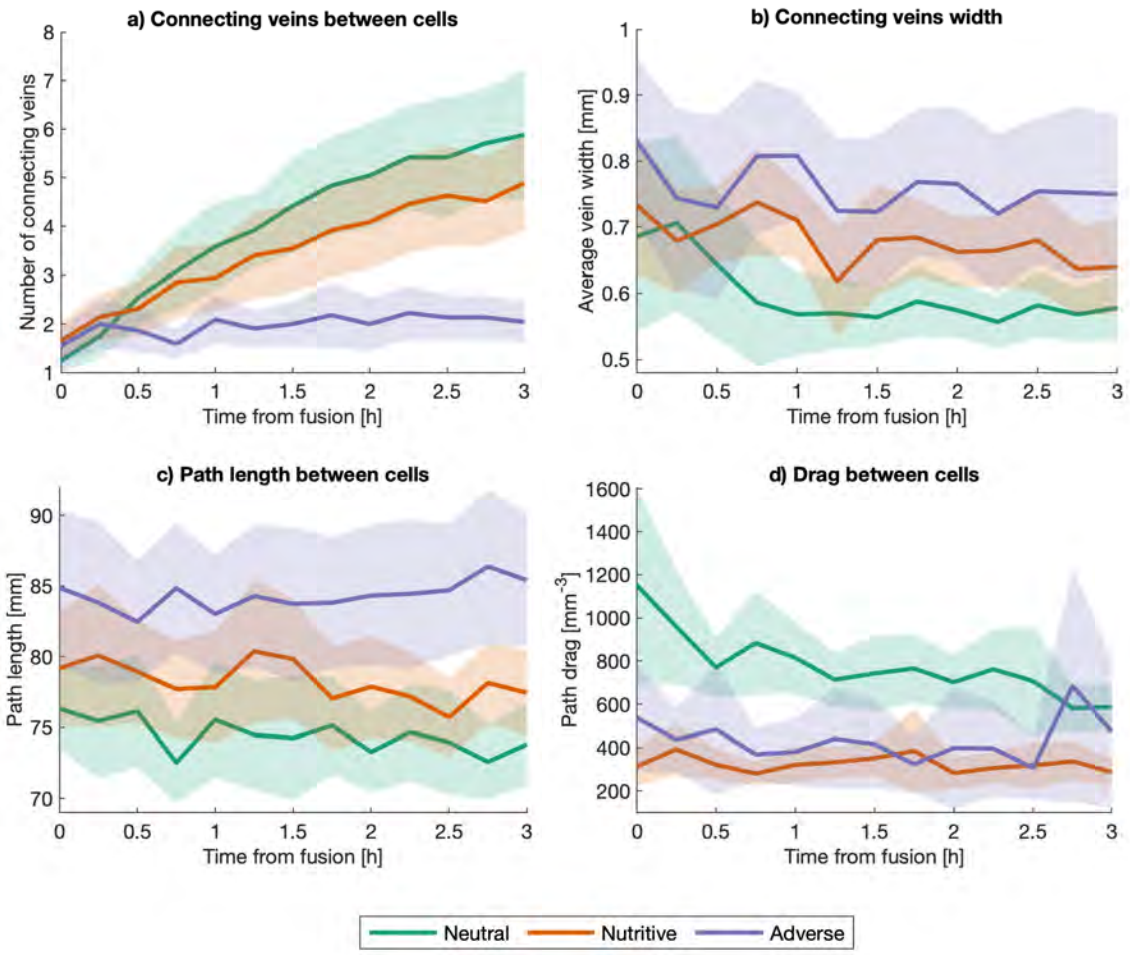

Figure 3.19: Connectivity between slime mold cells. a) Number of connecting veins between fused slime mold cells. b) Average width of the connecting veins, calculated as the total width of the region between slime mold cells, divided by the number of connecting veins. c) Path length between initial slime mold cells: shortest distance between the two initial slime mold cells along the network. d) Minimum drag along the path connecting the two initial slime mold cells in the network.

From Figure 3.19a, we can see that the number of connecting veins in the neutral and nutritive environments consistently increases after fusion, while it remains relatively constant in the adverse environment, with between 1 and 2 veins connecting both sides of the initial slime mold cells. The average width of the veins connecting the slime mold cells remains relatively constant after fusion (Figure 3.19b), with the networks in the adverse environment showing a higher connecting vein thickness than the ones in the neutral environment. These results suggest that slime molds in the neutral environment create several,

relatively thin connecting veins between slime mold cells, in contrast to the adverse environment, in which the connection between slime mold cells depends on few (around 1 or 2) thicker veins.

Moreover, the shortest path between the connected slime mold cells is significantly longer in adverse environments, as shown in Figure 3.19c, which might be a consequence of the growth dynamics of slime mold in such environments. A neutral environment fosters a 'spread-out' network, which is more likely to fuse with the other slime mold cell at mid-distance, eventually resulting in a shorter path between slime mold cells. On the contrary, the networks in adverse environments follow a more compact growth direction, taking significantly more time to fuse (Figure 3.15) and possibly resulting in fusing regions in more diverse locations, through which connecting veins develop, eventually resulting in longer connecting paths.

The drag between slime mold cells, shown in Figure 3.19d, shows that even though the path length between slime mold cells in the neutral environment is lower than that in the adverse environment, the fact that the veins in the neutral environment are thinner (including the connecting veins) results in a higher drag between slime mold cells. Moreover, while the drag and path lengths between slime mold cells remain relatively constant in the adverse environment, they decrease on the neutral substrate, suggesting the networks in the neutral environment further optimize the connection between slime mold cells after fusion.

3.5 Discussion

The analysis of slime mold networks and growth dynamics with a semi-automated image processing approach confirms that substrates have a direct impact on the growth rate of the cells [27]. Growth slows down in nutritive and adverse environments, which yields an earlier time to fusion in a neutral environment. Based on a comparison of slime mold cells of same area, it was found that slime molds in the neutral conditions follow an exploratory behavior, creating spread-out networks quickly, while slime molds in adverse environments

minimize their number of connections, extent and growth rate, to avoid exposure to the repellent substance, while still exploring the domain. On the contrary, slime molds in the nutritive environment seem to focus on metabolizing the glucose, pushing aside the exploratory behavior. This observation is supported by the total network length and average vein widths of the substrates, which shows that slime molds in neutral environments build longer networks with thin veins, as they are more 'spread-out', while nutritive and adverse substrates promote compact, denser networks. Hence, our findings confirm previous observations [138] that network characteristics depend on the environmental conditions and our results are in agreement with previous observations on vein width [137, 134].

The connectivity of the different networks is not affected by the substrate or time period of growth of the networks, and we observed a linear relationship between connectivity and the number of edges and nodes (number of edges around 1.35 times the number of edges), which supports the fact that the mean node degree of slime mold cells is a little below 3 – a common feature of biological networks [140, 136]. Moreover, when comparing the networks to the theoretical bounds, e.g. simply (MST) and fully (DT) connected networks, we found that slime molds yield total lengths closer to simply connected networks with low tortuosity, suggesting a cost minimization strategy.

Slime mold networks contains loops, even though slime mold seems to minimize its network length, as demonstrated by a relatively high transport efficiency when compared to the straight distance between nodes. Such efficiency proved to be higher in neutral and adverse environments when compared to nutritive environments. Conversely, the drag efficiency of the networks built in the neutral substrate was the lowest, which is congruent with the fact that neutral environments promote a spread-out network (longer distances), with thinner veins (higher drag). The normalized network length (relative to the MST and DT) proved to be proportional to the normalized fault tolerance (resiliency) of the networks, while it was inversely proportional to the mean drag betweenness (network centrality). These findings show that longer networks (with the same total area) are more connected

and therefore are more resilient and less centralized.

The similarity between the network dynamics of the networks built in neutral and adverse environments arises again when considering the load balance and resiliency of the networks. Slime molds in the nutritive environment produce more centralized networks, e.g. the traffic depends on few edges [140], which also explains the fact that such networks are less resilient, becoming more disconnected as some edges disappear from the network. This observation suggests that even though the slime molds in neutral and adverse environments have significantly different growth rates, they still follow a similar exploration strategy, while the slime molds in the nutritive environment follow a different dynamic due to the presence of the glucose being absorbed.

By observing the fusion process between slime mold cells, we noted that after coming into contact with each other, usually via two pseudopods, a dense, fusion region develops. And after a certain point (about 20 minutes), that region starts thinning at different rates (faster in the neutral environment), going from a dense region to an array of connecting veins, thinner in the neutral environment. In terms of connectivity between fused cells, the fast exploration dynamics of the slime molds in a neutral environment results in a highly connected network, with several, thin connecting veins between the previously disconnected cells. On the contrary, the networks in the adverse environment remain connected with few (1 or 2 in average) but thick veins. Regarding the path length and drag between the connected cells in a neutral environment, the number of connecting veins between the two slime mold cells results in shorter (more efficient) paths, compared to the other two treatments. And even though veins in the neutral environment are generally thinner, their thickness evolves to reduce the drag between the connecting paths.

Through our study, we show that slime molds build elaborate networks that are highly responsive to environmental conditions. These networks develop as the organism explores its environment for new resources. In many vascular networks observed in living systems, efficiency corresponds to a measure of how fast nutrients and oxygen can be transported

along the network, while cost measures how much energy or carbon is needed to construct such network. Slime mold networks, similarly to fungal networks, are built following local iterative expanding steps rather than a pre-planned blueprint. The growth involves building links and nodes in excess, followed by selective pruning and reinforcement of particular links. It has been demonstrated that coupling between strengthening and elimination in slime molds allows the network to transit from a fine mesh in a neutral environment to an optimal solution when food resources are added in the environment [153].

We observed that slime molds networks in exploration mode exhibit many loops. Loops are a common feature of natural networks such as as animal, plant and fungi vascular networks [154]. It has been suggested that the redundancy of vascular networks could be an adaptation to the varying physiological demands of different parts of the system and a way to withstand damages [155]. The existence of loops prevents disconnection of the network. In the absence of loop, severing a vein would result in the loss of all the network sections downstream from that vein. In contrast to plants and animals in which the vascular networks form only a small part of the organism and is usually protected and preserved from the environment, in both slime mold and fungi, the network defines the organism itself and must adapt constantly to variable environmental conditions. Hence, vascular networks in plants and animals are usually very efficient and low-cost, but unadaptable, while fungi and slime mold networks tend to be robust and efficient, but costly[156].

More broadly, deploying an efficient network in a constrained environment with a finite amount of resources is a common objective to both biological and human-made systems. Regardless of scale, ranging from ecosystems and communities, such as ant colonies, to simple organisms and organs, such as neural arbors and vascular systems, it has been observed that network deployment strategies seek to maximize travel efficiency and/or redundancy (and therefore resiliency) while minimizing cost. Commonalities between natural and engineered networks have originated the field of bio-inspiration, which aims to mimic biological principles to improve the design and construction of human-made systems. For

instance, the principles of networks scale have shown to be applicable to both biological, human, and geographic networks expanding the similarities between networks to the principles behind the formation of basins and hydrology systems [157]. Comparison between biological and computational networks showed the differences between sparse and robust networks in terms of connectivity and optimization strategies [158] and highlighted the potential to combine biological principles and engineering control to yield more efficient networks than those created with each strategy alone [159]. Other studies have focused on transportation networks. For instance, Patino-Ramirez et al. [26] showed that a bio-inspired road network design based on leaf venations can reduce the unitary construction cost of the networks, while yielding similar path efficiencies to those obtained with global optimum networks, such as Steiner Trees. Similarly, Tero et al. [25], showed that slime mold creates efficient algorithms in terms of cost, path efficiency and resiliency, that can outperform current design methods used to design railway systems.

Civil infrastructure networks are traditionally designed to optimize objective functions (such as traffic flow) under static constraints (such as land use regulations). Network resilience and security have recently gained attention in several fields of engineering that study the impact of global threats such as pandemics, terrorism or climate change [2]. It has become increasingly important to design networks that adapt to disruptions (e.g., cut connection after a natural disaster) or attacks (e.g., data hacking or spread of a disease). In this context, classical optimization under constraint is not applicable. Storm water management relies on the simulation of weather scenarios, in which the final design option, e.g. pipe upsizing, underground storage, or bio-filtration, remains empirical [160, 161]. Probabilistic approaches were used to assess the vulnerability of power systems [162] and virtual networks [163], such as internet and optical networks, to physical infrastructure damage. The most recent publications treat the resilience of inter-related networks such as power, water and cellular networks [164]. In all of these models, analyses are done *a posteriori*. More dynamic approaches based on the game theory were adopted to secure

infrastructure and information networks, whereby the cost of creating and removing links in a graph is calculated at every move of the attacker or the defender [165]. No strategy has been proposed yet to account for the splitting and fusion functions of slime mold networks, which relate to specialization vs. generalization behavior. How to split water networks to ensure autonomy of communities in remote areas? When should two countries merge electrical power grids after a hurricane? So far, these questions are open. To the authors' best knowledge, the only studies available to date focus on data fusion [166, 167, 168] or on emergency operation center fusion [169]. The dynamics of slime mold networks before and after fusion in response to various environmental constraints has a potential to inspire new strategies to design adaptable information and infrastructure networks, resilient to natural and biological hazards as well as geopolitical risks.

3.6 Concluding statement

slime molds are a good model organism to study networks. First, slime mold networks are both a transport system and the organism itself, and are therefore more dynamic than vascular networks in plants and animals. Second, slime mold networks adopt an array of topologies in response to environmental conditions. Third, slime mold networks can be continuously remodelled via pruning and reinforcement, adapting to different nutrient conditions and damage. Fourth, slime mold networks differ from other types of vascular networks, due to an unlimited capacity for expansion, combined with an ability to maintain its functionality as a living unit. Lastly, self-fusion yields optimised interconnected networks.

CHAPTER 4
SLIME MOLD-INSPIRED FLUID EXTRACTION MODEL FOR RESERVOIR ROCKS

4.1 Aim and summary

The current chapter shows and application of the inferred algorithms and growth strategies of slime molds. Based on a slime mold inspired flow model, we propose a framework to optimize the connecting path between a distributed resource inside a porous medium (rock), and a point/s of injection or extraction, hence maximizing the fluid flow to/from the extraction/injection point. The contents of this chapter has been published in [29].

This study is motivated by the similarities between biological and engineering flow systems, which maximize their efficiency following the path of minimum energy. Industrial fluid extraction and injection processes are designed to minimize the implementation cost (energy, materials) and maximize the volume of fluid injected (or withdrawn), given a spatial distribution of resources (or pore that contain these resources). We explain the commonalities between the equations governing flow in a porous medium and growth of slime mold, an organism that dynamically deploys tube-like structures and adapts them as a function of their contribution to the overall network. We perform several simulations to analyze the influence of the pore size distribution and of pore spatial distribution on the topology of the extraction network predicted by the slime mold growth algorithm. We discuss the suitability of the biomimicry model to design fracture patterns for optimal fluid extraction from a porous rock.

4.2 Introduction

What fracture network topology optimizes fluid flow in a rock mass? Power laws and fractal dimensions govern most physical processes observed in nature. However, optimization algorithms that shape natural self-organization remain in many cases unknown. Furthermore, engineering approaches rarely consider nature as a guide to improve human systems and very few attempts have been made to explain the fractal nature of the topology observed in many connectivity networks. The constructal theory [170] is an optimization method for finite flow systems with configuration (e.g., slenderness), in which a purpose (e.g., heat collection) and constraints to the flow (e.g. mass balance) are established. This theory can be used to predict the topology of both natural networks, such as the design of the lungs [171], and man-made networks, such as the layout of cities [172, 173]. Unlike other biological models of cell networks, the constructal theory is completely deterministic, predicting the layout of a network without prior assumptions. Major limitations are: (a) the need to assume controlling thermodynamic potentials, steady boundary conditions and uniform distribution of connecting points; (b) the absence of interaction dynamics from the formulation which disregards the fluctuations of the environment and the corresponding adaptation of the network. In the same way, Steiner tree algorithms allow finding: The minimum spanning tree from a set of points distributed in space; The global optimum from the creation of Steiner points that minimize the length needed to connect 3 points on the plane; The minimum network length concatenating those full Steiner trees using a combinatorial algorithm [174].

The astonishing similarities noted between the geometry of networks formed by living organisms (e.g. roots and bacteria) and that of infrastructure facilities (e.g. railway systems, see [84, 25], inspired the present study. We follow the growth principles of true slime mold (*Physarum polycephalum*), an intelligent organism that yields better solutions than Steiner trees for network configuration [175] and is capable of changing its topology over time,

depending on the variability in topological and environmental conditions. Using a slime mold growth algorithm, we calculate the size of each crack that connects pores in a rock mass in order to maximize the extraction of a fluid from those pores. We first explain the principle of the bio-mimicry model used for fluid flow. We then study homogeneous and heterogeneous rock microstructures, with different pore size distributions. Lastly, we recommend designs of fracture patterns for optimal extraction from porous rocks.

4.3 Biological analog of fluid flow

4.3.1 Governing equations of the bio mimicry model

We construct a biomimicry flow network model, in which the porous medium is represented by a distribution of pores (nodes) of various sizes. Cracks (edges) connect the pores and allow the flow of fluid from the pores (sources) to a sink point. The width of the cracks connecting pores is calculated based on a mathematical model that mimics the growth behavior of *Physarum plucephalum*. This amoeba-like organism is a single cell consisting on a dendritic network of tubular elements that ensure the flow of nutrients and signals through the plasmodium [88]. The organism spreads uniformly under constant environmental conditions and adapts dynamically due to the composition of the substrate, the supply of light, the distribution of food spots and the presence of physical barriers that restrict growth [176]. Physical observation shows that the placement of a food source on certain locations of the network induces tube shrinkage and thickening, since flux through a tube strengthens it, while unused tubes tend to degenerate and disappear. The mathematical formulation of the present model follows the empirical rules established by [177], where tubes that are not part of the flow network tend to disappear, and, when two or more paths connect the same set of food spots, only the shortest prevails.

Flow through tubes is assumed to follow the Hagen-Poiseulle equation, which states

that the flux along a given tube can be expressed as:

$$Q_{ij} = \frac{\pi r_{ij}^4}{8\eta} \cdot \frac{p_i - p_j}{L_{ij}} \tag{4.1}$$

Where η is the dynamic viscosity of the fluid (assumed to be water in the present study, for simplicity), p_i and p_j are the pressures at nodes i and j respectively, r_{ij} and L_{ij} are the radius and the length of the tube, respectively. The conductivity of the tube is defined as:

$$D_{ij} = \frac{\pi r_{ij}^4}{8\eta} \tag{4.2}$$

The conductivity of the tube (edge representing a crack) changes when the radius of the tube changes. The length remains constant, and the pressure at each node is computed from the principle of mass conservation, as follows:

$$\sum_{i}^{n} Q_{ij} = 0 \tag{4.3}$$

The flux is set as $-I_{ij}$ and I_{ij} for nodes set as sink and source, respectively. The value of I_{ij} is proportional to the volume of the pore (node).

The resulting linear system of equations is solved for p_i and Q_{ij}. The change of tube conductivity within a time step n is given by the expression shown in the following equation, taken from [177].

$$\frac{\partial D_{ij}}{\partial n} = \frac{\left(\frac{q_{ij}}{q_m}\right)^{\gamma}}{\left(\frac{q_{ij}}{q_m}\right)^{\gamma} + 1} \cdot \mu_{D_{ij}} - rD_{ij} \tag{4.4}$$

Where q_{ij} is the flux through the tube during step n, q_m is the mean flux trough the network, D_{ij} is the mean conductivity of the network. γ is the parameter that describes the non-linearity of the growth feedback and is set as 1.8 according to [25] and r is the decay rate of the tube (set as 2%).

4.3.2 Slime mold growth algorithm

The model consists of a two dimensional, regular, triangular lattice mesh: the coordination number of a node is equal to 6, i.e. each node is connected to 6 other nodes by 6 throats that are 60° apart from each other. The grid contains n_x by n_y nodes, where n_x and n_y are the numbers of nodes along the x and y-axis, respectively. Each node represents a spherical pore inside the rock mass. Each edge that connects two pores represents a pore throat, modeled as a slime mold tube, which is assumed to be a cylinder. We assume that pore and pore throats follow independent normal size distributions. The parameters of the initial probability density functions are taken from published articles about reservoir rocks. Node (pore) radius distributions do not vary during the simulations, but the radii of the cylindrical edges are optimized by using the slime mold growth algorithm.

After generating the mesh and assigning the initial size distributions to the nodes and edges, n nodes (pores) are selected to play the role of flow sources. The probability of choosing a pore is proportional to its size, for instance the probability of selecting a pore of radius $2r$ is twice the probability of choosing a pore of radius r. At each computational step, we chose n equal to $0.5n_xn_y$ sources, which yielded a realistic network topology within a simulation time shorter than a day. Larger sets of sources did not yield significant changes in the topology of the system.

At each step of the computation, different sets of sources are considered but the sink point remains at the same location, at the top center of the domain ($i = 1$ and $j = 0.5n_x$). Next, the corresponding flux input and output are set according to the volume of the source pore, and the node pressures and edge flux values are obtained from the corresponding set of linear equations. Then, the conductivity and radius of each edge are updated and a set of nodal sources is considered.

4.4 Bio inspired optimization of fluid extraction from porous rocks

4.4.1 Mesh and boundary conditions

The simulations presented in the following were all done with the same lattice mesh, with $n_x = 50$ and $n_y = 70$. The scaling factor of the mesh (which defines the length of each edge) was set as 1, i.e. the mesh is considered non-dimensional according to the growth mechanism defined by [177].

The microstructure parameters where chosen according to [178] for a sample of Lithic feldspar sandstone in a tight oil reservoir. The pore radius is distributed normally with a mean value $\mu_p = 145\mu m$ and a standard deviation $\sigma_p = 35\mu m$; the pore throat mean radius is $\mu_t = 0.48\mu m$ with a standard deviation of $\sigma_t = 0.12\mu m$.

4.4.2 Microstructure models

We design different microstructure models to evaluate the influence of the variability of the pore and pore throat size distributions on the deployment of the slime mold inspired flow network. We study: The influence of the mean value and standard deviation of the pore (node) size distribution; The influence of the spatial variability of the pore sizes.

Influence of pore mean size on network topology

The three micro structures described in Table 4.1 were compared to study the effect of the variation of the mean pore size distribution on the development of the flow network.

Table 4.1: Effect of pore size. Mean and standard deviation of pore sizes used in the models.

Model	Mean (μ) [ηm]	Std. Dev (σ) [ηm]
1	145	35
2	290	35
3	435	35

Influence of pore size variability on network topology

In order to assess the influence of the standard deviation of the pore size distribution, four microstructures, described in Table 4.2, were compared. The mean pore size was kept constant as well as the pore throat size distribution parameters.

Table 4.2: Effect of pore size variability. Mean and standard deviation of pore sizes used in the models.

Scenario	Mean (μ) [ηm]	COV [%]	Std. Dev (σ) [ηm]
1	145	5	7.25
2 (base)	145	24.14	35.00
3	145	50	72.50
4	145	100	145.00

Influence of pore size spatial distribution on network topology

In order to understand the influence of the spatial distribution of pores on the deployment of the network, three different cases were considered, as explained in Table 4.3 and Figure 4.1. In each case, the domain is made of a combination of regions of different mean pore size, but same standard deviation. In case A, the whole domain has a low mean pore size. In case B, the lower half of the domain has a mean pore size three times higher than the upper half. In case C, the mean pore size is higher in the upper half than the lower half. The position of the sink is the same in all three cases: at the top center of the domain.

Table 4.3: Pore size distribution parameters. Mean and standard deviation of pore sizes used.

Pore Size Distribution	Mean (μ) [ηm]	Std. Dev (σ) [ηm]
PSD-1	145	35
PSD-2	435	35

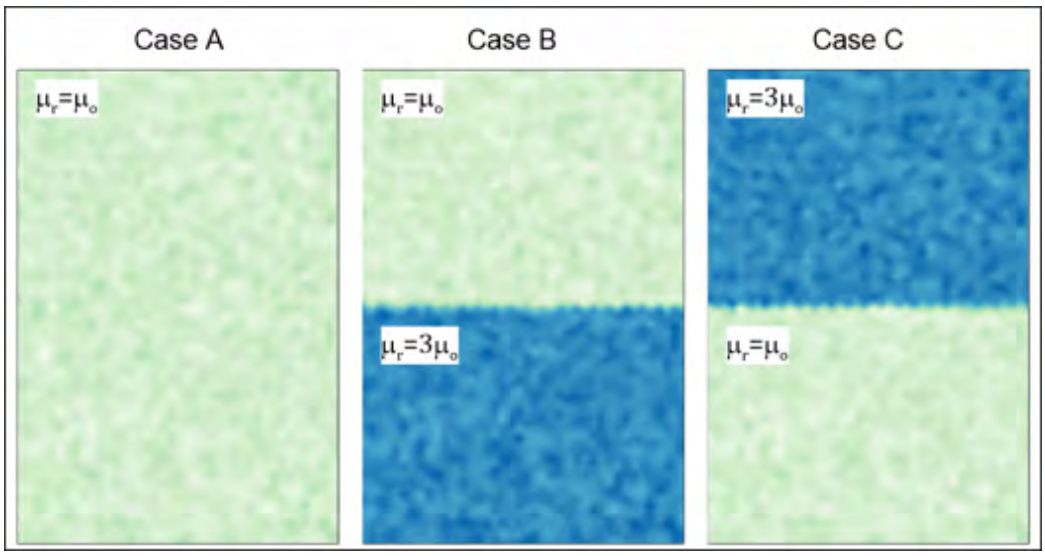

Figure 4.1: Pore size distributions for different cases studied.

4.5 Results

We compare the topology of the initial porous network (with pores and pore throats of given initial size distributions placed on the grid) to that of the final network, obtained after using the slime mold growth algorithm. Within the final network, we define a preferred flow path, as follows. The 10% larger edges obtained at the end of the simulation are selected (note that larger edges are tubes that are were selected the highest number of times to ensure flow from a set of pore source to the sink). Within the set of 10% larger edges, clusters of connected tubes are identified. In all simulations, one cluster contained 90% of the larger edges. This cluster defines the preferred flow path.

4.5.1 Influence of pore mean size on network topology

The average growth of tube (edge) radius was computed for edges within the entire domain and within the preferred path only. Results are summarized in Table 4.4. The growth rate of the edges within the preferred path is about five times higher than the average growth rate within the entire domain.

Table 4.4: Mean increase of edge radii with varying mean pore size.

Model	Every edge	Preferred path edges
1	3.42%	16.39%
2	3.52%	16.92%
3	3.39%	16.30%

The results show no significant differences between the increase of edges radii; this fact can be explained by the nature of the algorithm, which is sensitive only to the relative weight (size) the source points (pores) over the domain. Since the probability density function and standard deviations are the same in the three microstructure models tested, the variability of mean pore size does not affect the topology of the flow network generated with the slime mold growth algorithm. We check that the ratio between the total volume of the pores connected by the preferred path and the total pore volume is the same in the three simulations (Table 4.5). This result confirms that the same volume fraction of resources can be extracted through the preferred flow path in all cases.

Table 4.5: Effect of mean pore size on preferred path efficiency.

Mean pore Size	μo	$2\mu o$	$3\mu o$
% of pore volume connected by the preferred path	27.07%	26.59%	27.13%

4.5.2 Influence of pore size variability on network topology

Table 4.6 summarizes the edge growth rates within the entire domain and within the preferred path, for microstructures with different pore size standard deviations (described in Table 4.2). Table 4.7 shows the porous volume accessible by the preferred path. We note only slight differences in the four simulation results. This is because pore sizes are centered around the mean pore size. The probability of selecting a pore in a set is not strongly affected by the variability of the pore size distribution. Therefore the edges connecting the

pores are used with similar probability regardless of pore size standard deviation, which leads to similar network topologies.

Table 4.6: Mean increase of edge radii as a function of pore size standard deviation.

COV of pore size distribution	Every edge	Preferred path edges
5%	3.38%	15.90%
24.14%	3.50%	17.08%
50%	3.30%	15.60%
100%	3.94%	18.64%

Table 4.7: Effect of pore size standard deviation on preferred path efficiency.

COV	5%	24.1%	50%	100%
% Of total pore volume connected by the preferred path	26.75%	26.22%	24.98%	26.07%

4.5.3 Influence of pore size spatial distribution on network topology

The final networks obtained for the three microstructures described in Table 4.3 and Figure 4.1 (Cases A, B and C) are shown in Figure 4.2. The creation of preferential paths for fluid extraction is analyzed in terms of relative growth or shrinkage of the edges (percentage of variation of radius size). Results are plotted as heat maps where the color intensity corresponds to the increase in diameter of a given edge.

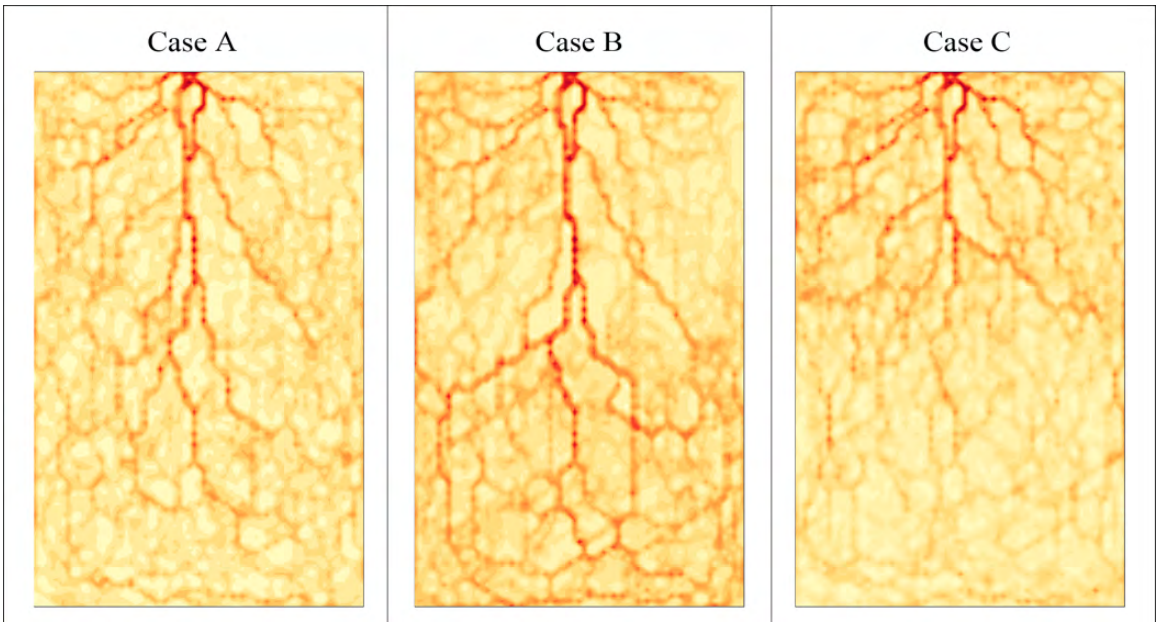

Figure 4.2: Topologies obtained for each case studied (A, B and C). The intensity of the color is related to the relative growth of the given edge within the domain defining the preferred flow path.

Table 4.8, Table 4.9 and Table 4.10 summarize the analysis of network connectivity, i.e. the average radius growth within the entire domain and within the preferred path only.

Table 4.8: Mean increase of edge radii – Case A

Mean increase of edge radii	Every edge	Preferred path edges
Entire Domain	3.14%	14.73%
Upper Half	3.54%	15.86%
Lower Half	2.73%	13.22%

Table 4.9: Mean increase of edge radii – Case B

Mean increase of edge radii	Every edge	Preferred path edges
Entire Domain	6.78%	33.32%
Upper Half	6.79%	36.86%
Lower Half	6.76%	30.08%

Table 4.10: Mean increase of edge radii – Case C

Mean increase of edge radii	Every edge	Preferred path edges
Entire Domain	1.51%	6.93%
Upper Half	1.79%	7.23%
Lower Half	1.23%	6.20%

As expected, the increase of the edge radii is significantly higher on the preferred path compared to the complete set of edges (around 5 times higher). Within both the entire domain and the preferred path, the highest increase is seen in Case B, where the higher concentration of pore (and resource) volume is far away from the sink. Tube growth is needed to increase conductivity and reach the resources in the lower part of the domain. Case C presents the smaller increase of edge radii overall, which can be attributed to the spatial distribution, where the biggest concentration of the resources is already close to the sink. There is no significant incentive to enhance flow far away from the sink according the slime mold growth algorithm. Finally, we note an important difference of growth rate between zones: the growth rate is always greater in the upper half compared to the lower one; this result can be explained by the necessity to deploy a network of low resistance to the flow at the vicinity of the sink. In Table 4.11, Table 4.12 and Table 4.13 the first column shows the distribution of the pore volume in the domain. By definition, the sum

of the pore volume fractions in the upper and lower halves equals the pore volume fraction in the entire domain, i.e. 100%. In the same way, the sum of the volume fraction of pores connected by the preferred path in the upper and lower halves equals the volume fraction of pores connected by the preferred path over the entire domain. The efficiency is defined as the ratio between the pore volume connected by the preferred path and the total pore volume available in a given area.

Table 4.11: Preferred path efficiency - Case A

Location	Fraction of total pore volume	Fraction of total pore volume connected by the preferred path	Efficiency
Full domain	100%	27.01%	27.01%
Upper half	50.01%	15.59%	31.17%
Lower half	49.99%	11.42%	22.86%

Table 4.12: Preferred path efficiency - Case B

Location	Fraction of total pore volume	Fraction of total pore volume connected by the preferred path	Efficiency
Full domain	100%	26.83%	26.83%
Upper half	11.12%	3.00%	26.98%
Lower half	88.88%	23.83%	26.81%

Table 4.13: Preferred path efficiency - Case C

Location	Fraction of total pore volume	Fraction of total pore volume connected by the preferred path	Efficiency
Full domain	100%	32.84%	34.84%
Upper half	88.90%	31.16%	35.05%
Lower half	11.10%	1.68%	15.12%

Results show that with less than 10% of the edges of the domain, the computed network is able to reach around 30% of the total pore volume available, with similar results even for strongly heterogeneous distributions of total pore volume. Furthermore, the results show that the highest value of efficiency is obtained for Case C, where the resources are closer to the sink, similarly to the results obtained for mean increase of edge radii. On the other hand, the lowest efficiency (26.83%) is obtained in Case B, where the majority of the resources are away from the sink, which also agrees with the results obtained for edge radii increase. Once again it was observed that the efficiency is always higher in the upper half of the domain, where at the same time the preferred path shows highest radii increase as well. These results show the importance of network topology for efficient extraction of resources from porous rocks, through main veins connected directly to the sink.

4.6 Conclusions

We construct a slime-mold-inspired flow network model, in which the porous medium is represented by a distribution of pores (nodes) of various sizes. Pores represent source points, and a sink point is placed at the top center point of the domain. The size of the cracks (edges) connecting pores is calculated based on a mathematical model that mimics the growth behavior of *Physarum polycephalum*. The algorithm selects flow paths that preferably connect the larger pores. Edges along these flow path grow in size as a function of the number of times they are selected for conveying fluid flow from the pores to the sink. We compare the topologies of networks generated with different pore size distributions. In all cases, a unique preferred flow path can be identified. The top 10% larger edges connect pores that represent 25% to 32% of the total pore volume. The growth rate of edges within the preferred path is five times higher than the average growth rate of the edges in the entire domain of study. Edges tend to grow in size when they are closer to the sink. The slime mold inspired porous networks maximize fluid extraction. The results obtained show the potential to model real distributions of resources within porous media, in order to find

the optimal flow path based on slime mold growth; for this reason, further research on this topic will include benchmarking between fracture patterns observed in the field versus the bio-inspired network, in order to validate the effectiveness of the method. Lastly, it is important to mention that specific calibration for this application is needed, since the growth law adopted in the present article is directly taken from a biological model. Even though this choice is qualitatively suitable for the analysis, further research will include the development of physical experiments on slime mold.

CHAPTER 5
TRANSPORTATION NETWORKS INSPIRED BY LEAF VENATION ALGORITHMS

5.1 Aim and summary

The current chapter further explores applications of bio-inspiration to engineered transport networks. Specifically, this study evaluates the application of a leaf venation algorithm to transportation networks in the city of Atlanta. First, we benchmark leaf venation algorithms against theoretical bounds, and we then apply these algorithms to two proofs-of-concept models of the city, where the network follows the population distribution. The contents of this chapter has been published in [26].

In this study, we seek to apply optimization patterns exhibited by biological systems, which have adapted to environmental constraints and limited resource availability, to the optimization goals of human-made networks. Here, we evaluate the algorithm underlying leaf venation (LV) deployment using graph theory. We compare the traffic balance, travel and cost efficiency of simply-connected LV networks to those of the fan tree and of the spanning tree. We use a Pareto front to show that the total length of leaf venations is close to optimal. Then we apply the LV algorithm to design transportation networks in the city of Atlanta. Results show that leaf-inspired models can perform similarly or better than computer-intensive optimization algorithms in terms of network cost and service performance, which could facilitate the design of engineering transportation networks.

5.2 Introduction

The seemingly simple problem of connecting a central node to a set of spatially scattered points is common to many natural and artificial systems of all scales and levels of

complexity. Despite the differences in the mechanisms that drive network deployment, optimality is always sought as the maximum of objective functions under environmental constraints. Performance metrics are defined a priori by the modeler. A complex system can be optimized locally or globally, and optima may vary over time due to environmental constraints or internal changes such as ageing, growth and demand and supply. Complex optimization problems have been addressed by nature for millions of years by means of evolution; external stimuli trigger adaptation of organisms (and their organs) to their environment (constraints) and the competition for resources results in increasingly efficient organs, organisms, communities and ecosystems over subsequent generations [179, 180, 181, 182]. In closed environments with a limited availability of resources, biological networks constantly adapt to improve their efficiency, robustness or flexibility. Principles of thermodynamics impose some trade-offs: increased efficiency towards a specific function decreases the performance of the network in some other manner. In the case of fully connected graphs, minimization of the total network length results in a simply connected graph, having exactly one unique path between every pair of nodes [183]. Conversely, dynamic environments promote the development of robust and resilient networks, capable of overcoming accidents and errors and that have a better evolvability, at the expense of an increased network cost [184, 185, 186, 187, 188].

Organs and tissues develop strategies to optimize flow, energy, cost, connectivity and any other characteristic pertinent to their function. For instance, circulatory vessels are shaped to ensure efficiency and resilience [189], neural arbors deploy with minimum distance between cell bodies and synaptic partners and minimum network cost [190], plants optimize flow from roots to leaves while minimizing the total energy cost of their growth [191], root systems follow high hydraulic gradients [192, 179, 180], and uptake water as they grow, which makes the soil-root dynamics highly coupled and non-linear [193, 194, 195]. Remarkable optimization strategies were also noted in relatively low complexity organisms, including fungus [196], and slime mold, a unicellular organism capable of

achieving continuous optimization of its foraging path [175, 108]. Computational algorithms inspired by slime molds were used to aid solving NP-complete problems, such as finding Steiner trees [175, 176], and to assist the design of transportation networks [25, 177]. Communities have been studied extensively through ant colonies, which minimize the local cost of their networks [197], to the expense of reduced robustness [198]. Related research focuses on colony organization [199, 200], robustness versus efficiency [201] and computational optimization algorithms [202]. Other animals exhibit intrinsic optimization strategies [203, 204] in their travel and exploration dynamics.

Bio inspiration has been adopted in many disciplines as a vector of innovation. For instance, doctors practice surgeries with a mosquito-inspired needle [205], cod glycoproteins are used in the industry for their antifreeze properties [206] wind turbines were optimized by taking inspiration from the flippers of humpback whales [207] and mussels inspired the fabrication of novel adhesive compounds [208]. Engineering and natural networks present similarities both in their global and local optimization objectives. For instance, water networks were designed by solving an NP-complete problem using particle-swarm optimization, with a bird-flock inspired algorithm [209]. Internet networks were optimized with an algorithm inspired by slime molds [210]. Bus routes were calculated based on ants behavior [211].

In this study, we evaluate a leaf venation (LV) algorithm for designing bio-inspired infrastructure networks. Then, we describe an algorithm that models the mechanisms that drive LV deployment and we present two reference algorithms for benchmarking. In section 5.3 we compare the LV algorithm to the two reference models by means of a set of topological indexes and a Pareto optimality front. In section 5.4, we apply the LV algorithm to design a transportation network in the metropolitan area of the city of Atlanta (GA) and to expand the current metropolitan network towards an adjacent suburban county. And then, we present our conclusions regarding the potential of using LV algorithms for infrastructure design.

5.3 Network algorithms

LVs connect the stem to points distributed on the blade of a leaf with minimum length to transport fluids under environmental constraints. Similarly, transportation networks connect areas with high population or high economic activity and are designed to minimize cost and maximize efficiency during construction and service life. Based on this analogy, we propose to apply an LV algorithm to design a transport infrastructure network under the following assumptions: (i) The existence of a single source/sink point connected to an arbitrary number of attraction/service points; (ii) The absence of healing mechanisms that can overcome possible disconnections; (iii) The possibility of discretizing the domain into nodes that need to be connected (called attraction points from now on). In this section, we present leaf venations as simply connected networks embedded in a homogeneous space at steady state and we compare the performance of the LV algorithm to that of two benchmark network algorithms.

5.3.1 Leaf venation (LV)

Plant leaves grow a vascular system of interconnected veins, which ensure evapotranspiration and mechanical stability [180, 212]. Leaf venation systems form hierarchical networks, usually starting with a main single vein that grows from the petiole, followed by secondary branches that start from the main vein, and are connected at the same time to smaller veins. The latter, called tertiary veins, create paths that connect every single stoma to primary and secondary veins and therefore permit flow from/to the petiole to the whole leaf blade [213]. Additionally, tertiary veins form loops in the network, adding redundancy to the system so that in case a vein is cut, flow paths still exist to reach the entirety of the leaf blade; the existence of loops in the network increases its overall cost [155]. Here, we study cost minimization and thus focus on primary and secondary veins. We do not study the mechanical stability of the leaf and focus on the mass transport function of the LV.

Vein development [214] influences the shape and growth of the leaves [215, 216]. Leaf blade shapes are either simple (with a single unit) or compound (with two or more leaflets). Simple leaves can be entire, if they have a smooth or slightly toothed edge, or lobed, if they have significant indentations that make the contour highly non-convex [216]. Marginal growth characterizes leaves that develop outside of the current blade contour whereas diffuse growth refers to blades that stretch themselves to increase area [214]. In our study, we focus on simple, entire leaves that experience no growth; thus, we study the case of a relatively unconstrained geometric domain that grows outwards (no stretching of existing veins). This scenario is congruent with transportation infrastructure networks, which are constructed progressively, extending from the previous step of the network.

The most accepted algorithm for modelling vein patterns formation is based on the canalization hypothesis [217], which states that the growth and branching of new veins are controlled by the spatial distribution of a signal distributed along the leaf blade. This signal is in large part attributed to a growth hormone called auxin. Physical evidence shows that auxin sources can be viewed as attraction points discretely distributed throughout the blade [218], and numerical LV algorithms based on that assumption were validated against biological experiments [216]. The diameter of the veins obeys Murrays law, a power law with exponents that depend on plant species [219].

In the present study, we adopt the algorithm proposed by Runions [216]. The algorithm is described in Figure 5.1. In this implementation, the domain and auxin points (attraction points) are fixed at the start of the algorithm, and no new auxin point is added over time. Rectangular domains were used, and the kill distance was set as 0.5% of the largest dimension of the domain. After the LV architecture is obtained, it is transformed into an undirected graph, preserving the source, branching and attraction points and their connectivity, therefore making the edges between nodes straight lines.

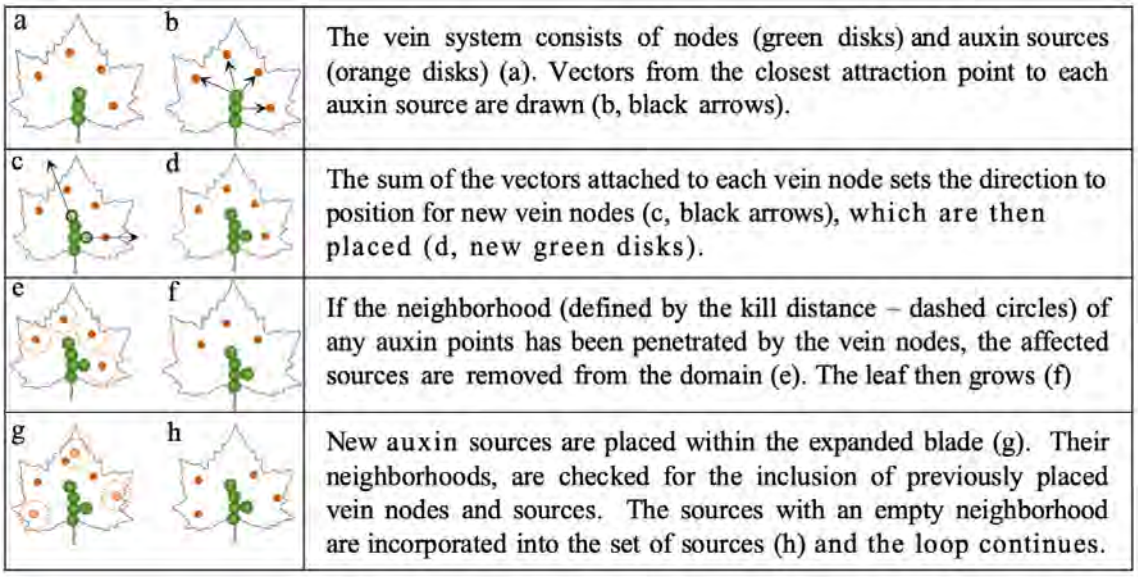

Figure 5.1: Leaf Venation (LV) pseudo-algorithm, after Runions [216].

5.3.2 Benchmark networks

In the following, let T_l be the network that minimizes each objective function. We benchmark the LV algorithm against a local optimization algorithm (fan tree), for which T_l is defined as the sum of the minimum source-attraction point distances, and a global optimization algorithm (Steiner tree), for which T_l is the minimum total length of the edges in the network. The fan and Steiner trees are constructed to connect a given set of points regardless of whether they are source or attraction points. The fan tree and the Steiner tree results are then used to construct a Pareto optimality front [220].

Fan Network

The total length of a network that minimizes each travel distance from the source node to an attraction point expressed as:

$$T_l = \sum_{i}^{n} d_i \tag{5.1}$$

Where d_i represents the distance along the edges from the source point to the i-th at-

traction point, $i \epsilon \{1, ..., n\}$.

The network that minimizes T_l is a fan network (FN), i.e. a collection of straight lines from the source to each sink. The FN represents a local optimum for each travel distance from the source. The nodes of the FN are only the source and the attraction point, i.e. there is no branching node.

Steiner Tree

A Steiner tree (ST) is a network that connects a set of points with the minimum total network length [221, 222, 174]. Additional branching nodes (called Steiner points) can be introduced in the system. The global optimization criterion for STs is as follows (Eq Equation 5.2).

$$T_l = min \left[\sum_{i}^{m} e_i \right] \tag{5.2}$$

Where e_i represents the length of edge $i \epsilon \{1, ..., m\}$. Note that m designates the number of node-to-node segments in the Steiner tree, which comprises nodes other than the source or the attraction points (Steiner points).

The basic Euclidean ST problem is in an unconstrained domain without obstacles [174]. Subsequent advanced algorithms that proposed to include obstacles and other geometries [223, 224, 225, 226] are NP-complete problems, i.e., the running time of the algorithm grows exponentially with the number of nodes [227, 228]. In order to circumvent the exponential increase of the runtime with the number of nodes, heuristics and alternative algorithms have been developed to approximate the solution and/or obtain an initial guess of the solution; some of these approaches have taken inspiration from bio-inspired systems such as slime mold growth algorithms [175, 176]. In the current implementation, we follow the algorithm proposed by Fonseca and collaborators [229] to find the Steiner trees, which are then transformed into undirected graphs for analysis.

5.3.3 Pareto Efficiency

We analyze the optimality of networks to satisfy: (i) A local criterion, to minimize the sum of the travel distances from the source/sink point to each attraction point along the edges of the network; (ii) A global criterion, to minimize the total network length. The Pareto optimality front [220] is an optimality line that indicates the smallest sum of individual distances that can be obtained for a given total network length.

In order to generate the Pareto front, we create solutions that follow a joint optimization objective, which is a linear combination between the two optimization criteria evaluated. The combination of the objectives is controlled by a parameter $0 \leq \alpha \leq 1$; when $\alpha = 0$, the objective reduces to the local criterion (Fan Network), while when $\alpha = 1$, the objective reduces to the global criterion (Steiner Tree).

To do so, we use the available plant-inspired greedy algorithm developed and implemented by Conn and collaborators [191], which constructs near-optimal architectures as shown in [190, 191, 230]. The algorithm starts from a stem protruding from the source node towards the centroid of the attraction points. From there, new branching points (and branches) are iteratively tested, choosing the branch that minimally increases the value of the objective (a function of α).

A network is considered Pareto optimal for the criteria tested if it lies along the Pareto front. We hypothesize that LV networks optimize one of the optimization criteria tested or a combination of both (along the Pareto front). Nevertheless, LVs are not only optimized for travel distances and cost, but also for other criteria not studied here [191, 220] related to mechanical stability, genetics, heat transfer among others.

5.4 Evaluation of optimality

5.4.1 Arbitrary networks evaluation

In order to analyze the underlying optimization mechanisms of leaf-inspired algorithms, we generated networks that connect a source to randomly distributed sets of attraction points in a 2D rectangular domain. The source node (petiole of the leaf) was placed at the bottom center of the domain (coordinates [0,0]). We tested 50 replicates of 10 attraction points (auxin points), and 50 replicates of 15 attraction points. The coordinates of the attraction points were uniformly distributed, in the range [-50 50] along the X-axis, and in the range [30 200] along the Y-axis. The Y range was set to start from 30, to create a distance between the source node ([0,0]) and the rest of the nodes, and therefore, enhance the tree-like structure of the resulting networks. The networks obtained are characterized as undirected graphs, in which the nodes correspond to the input points plus the set of branching points generated by each one of the algorithms. Figure 5.2 shows an example of the obtained networks for a set of 10 attraction points.

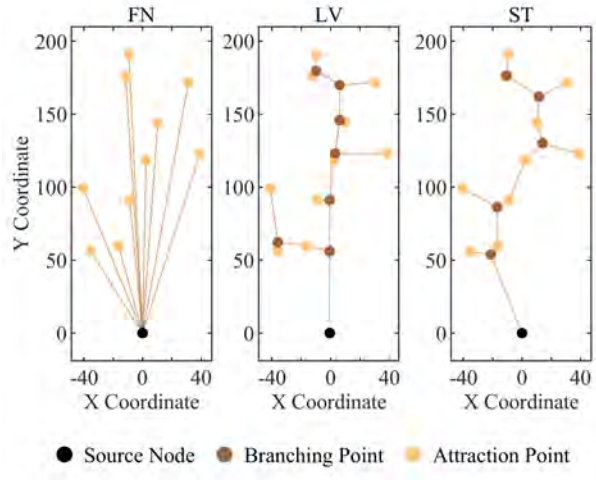

Figure 5.2: Example network topologies: Fan Network (FN), Leaf Venation (LV) and Steiner Tree (ST). Results for a set of 10 random attraction points.

5.4.2 Pareto Optimality

For each of the 100 sets of attraction points (called replicates in the following), we generate the LV, the FN, and the ST. For each network, the sum of individual travel distances from the source to each attraction point along the edges of the graph (local index) and the total network length (global index) are calculated. To aid visualization, we normalize the indexes according to the optimal bounds. The FN exhibits the lowest sum of individual travel distances and therefore its local index is mapped 0 and its global index set as 1; conversely, the ST exhibits the lowest total network length and therefore its global index is mapped to 0 and its local index set to 1. Then, the indexes of the LV are mapped based on the global and local bounds. Figure 5.3 shows the results. The shaded region encloses the Pareto fronts of all the replicates, which are also normalized. Contiguous boxplots show the variability of the sum of the individual travel distances and of the total network length among the 100 LV networks.

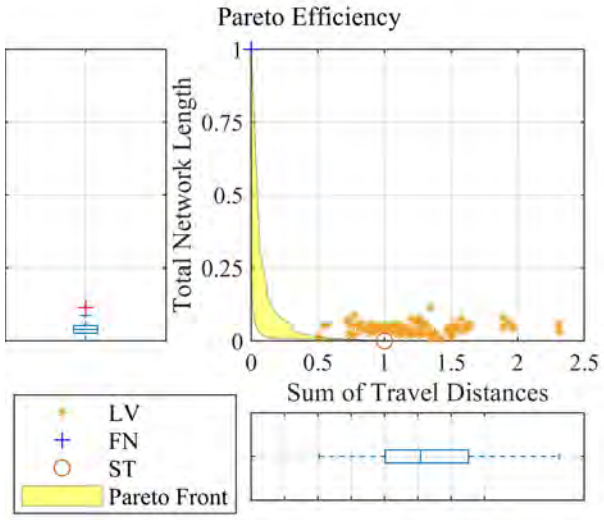

Figure 5.3: Normalized Pareto front, for 100 replicates of randomly distributed attraction points. Scattered dots correspond to the results obtained with the LV algorithm. Adjacent boxplots illustrate the variability of the local and global network indexes.

The total network length of the LV networks is close to the theoretical minimum (median value of 0.04) with a small variability (maximum value 0.11 and minimum value of

0.01). On the other hand, the sum of travel distances from the source shows high variability, ranging from 0.51 to 2.31 with a median of 1.27. This result suggests that the LV algorithm seeks to minimize the global network length but does not optimize the sum of individual distances to the source.

5.4.3 Service life performance

The optimization criteria used to plot the Pareto front are related to the initial design of the network. In addition to cost, travel path efficiency and load (traffic) balance are also important performance metrics of a transport network during its service life. In the following section we study the service life performance assuming that the capacity (proportional to the width of the edges) is homogeneous and fixed for all the edges of the graphs.

Load Balance: Edge congestion

Load or traffic balance is a common measure of service performance that has attracted a lot of attention in the field of graphs theory [183]. Depending on the context, the interpretation of load distribution has different meanings:

Congestion: if the capacity or width of the nodes or edges of the network are fixed and/or homogeneous, the system may suffer from bottleneck congestion, where the transport efficiency is controlled by the most congested link; therefore a uniform load distribution is optimal – like in parallel computing [231].

Centrality: if the capacity of the elements of the network is not a limitation, the distribution of the traffic load is a measure of the node/edge importance inside the network; therefore an irregular distribution of load is desired in order to classify or identify components in the network – like in social media networks [232, 233].

A common metric to find the load distribution in a graph is the node betweenness centrality [232]; it is defined as the number of times a given node is part of the shortest path between two other nodes of the network, normalized by the number of nodes in the graph.

In the current analysis, we use a slightly modified metric of load balance: we compute the load in the edges rather than the nodes, and we consider the traffic between every pairwise combination from the set of source and attraction nodes (excluding the branching/Steiner points). That way, the load balance only stems from the set of nodes that is common to all networks. For each of the algorithms evaluated in the benchmark, we plot the mean value of load balance and the interquartile range (IQR), a measure of dispersion that corresponds to the difference between the 25th and 75th percentiles of the data. Results are shown in Figure 5.4.

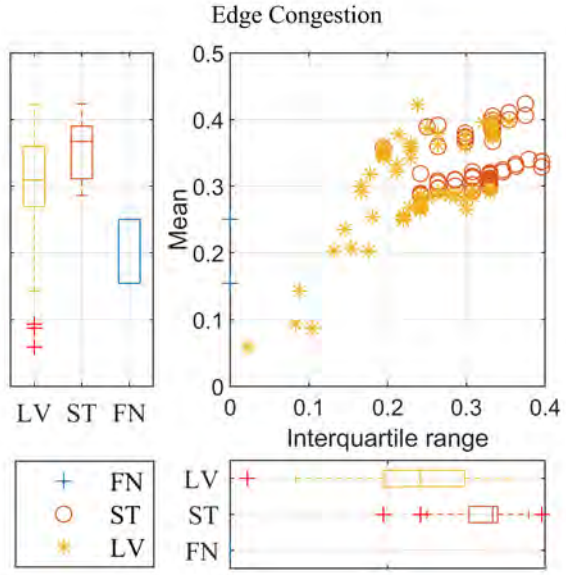

Figure 5.4: Edge congestion distribution. Vertical values show the mean edge congestion for the different networks while the horizontal axis shows its variability, in terms of the interquartile range..

For the current assumption of homogeneous and fixed edge capacity, we find that the LV and ST solutions have similar values of edge congestion, with median mean values of 0.31 and 0.37 respectively. Nevertheless, the traffic load is more evenly distributed in LVs than in STs: the IQR medians are 0.24 and 0.33 for LV and ST, respectively. On the other hand, the FNs outperform the STs and LVs, with a smaller median mean load (0.25) and a perfect load balance (IQR=0) - since all the edges are used the same number of times.

These results suggest that longer networks are more likely to balance traffic, since resources are not limited. But as the total network length decreases, the load is concentrated in certain edges, causing increased edge congestion.

Travel path efficiency

Besides load balance, the efficiency of traffic or flow during service life depends on the travel distance between pairs of nodes. We define travel path efficiency as the ratio between the travel distance between two nodes in a network and their Euclidean distance in the 2D domain. We first evaluate the paths from the source node to each attraction point, as shown in Figure 5.5. We then analyze the path efficiency between every pair of nodes (excluding branching points), as shown in Figure 5.6.

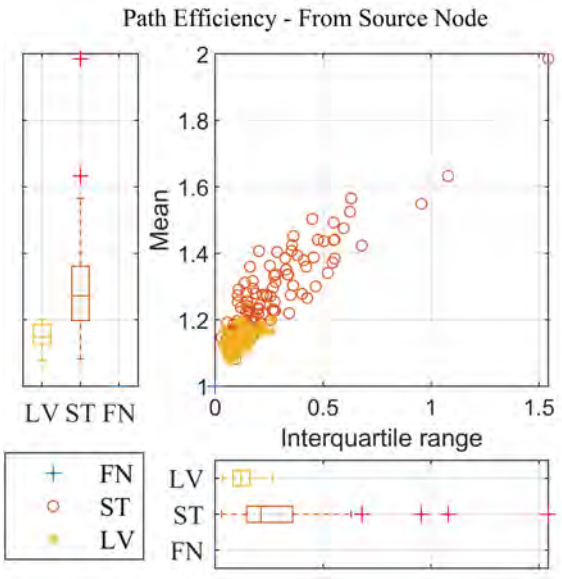

Figure 5.5: Travel path efficiency, paths from the source node to the attraction points.

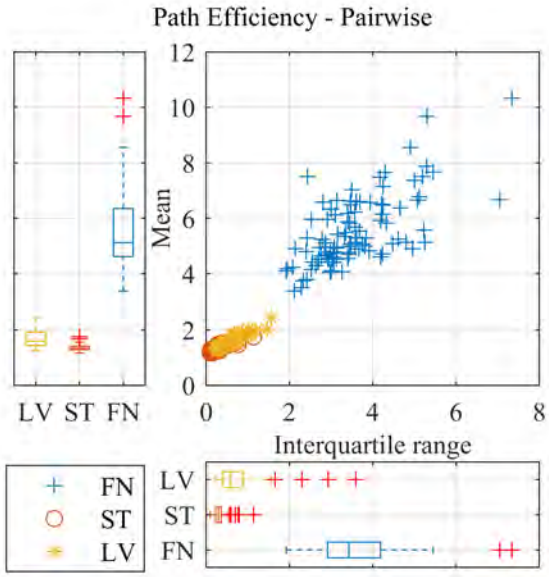

Figure 5.6: Travel path efficiency, paths joining pairs of attraction points.

By definition, the FN algorithm exhibits perfect path efficiency from the source, where every path has an efficiency of 1 (no dispersion). The LV algorithm outperforms ST both in terms of mean value and dispersion: LV median mean efficiency and IQR are 1.15 and 0.12, against 1.27 and 0.21 for the STs. Path efficiency among pairs of nodes cannot be optimal for simply connected graphs because edges are used in multiple paths. Not surprisingly, FNs have the lowest performance in this index, with median mean values from 3.38 to 10.34, because the load has to travel through the source to reach the destination node. The ST algorithm outperforms LV: the median mean value is 1.59 for LVs and 1.34 for STs, and the median IQR is 0.59 for LVs and 0.29 for STs.

5.5 Proof of concept: transportation networks in Atlanta, GA

Urban transportation networks are designed at minimum length (or cost) for optimal path efficiency and traffic balance, under land use and budget constraints. Here, we compare the performance of the LV algorithm to that of the ST for designing transportation networks in the city of Atlanta (GA), for which the population density map is available from the census

of 2010 [234].

5.5.1 Atlanta metropolitan area

We start by studying the five most populated counties of the metropolitan area of the city (Fulton, Gwinnett, Cobb, DeKalb and Clayton), and we compare LV and ST networks with a uniform-fixed edge capacity. The metropolitan area of Atlanta spans radially from an economic and geographic center. The natural center of the city (and the most densely populated area) is downtown, which also concentrates a large amount of venues of interest including auditoriums, stadiums, touristic sites and business and commercial operations. Therefore, we set the source node of the network as the location of the actual hub on the railway lines: Five Points station in the heart of downtown.

The attractions points are population centroids. We use a weighted k-means algorithm [97] to calculate the position of fifteen attraction points. Each attraction point is the weighted centroid of the region that contains the population that lives closer to that attraction point than to any other attraction point. Figure 5.7 shows the distribution of the density of population by census tract, shown as shaded regions, and the attraction points. For context, we also show the actual metro system of the city: the MARTA railway. The LV and ST networks are shown in Figure 5.8 and Figure 5.9, respectively. The FN was computed as well to calculate the Pareto optimality line.

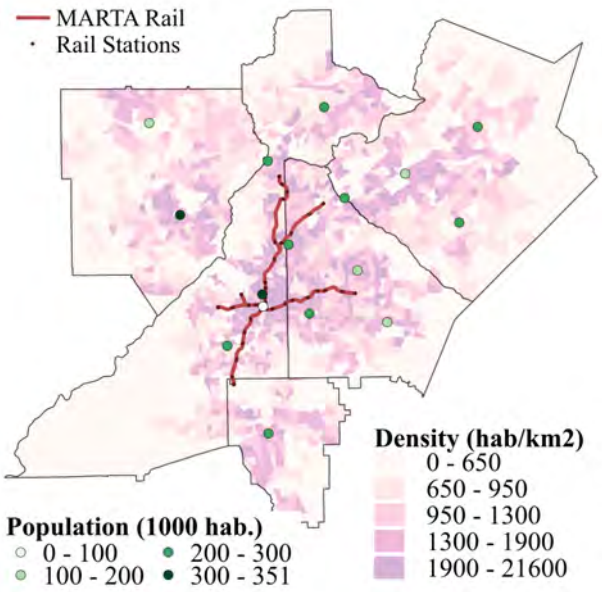

Figure 5.7: Density of population of the metropolitan area of the city of Atlanta (shaded background) and the corresponding fifteen population centroids (red dots). MARTA railways network shown in red solid lines.

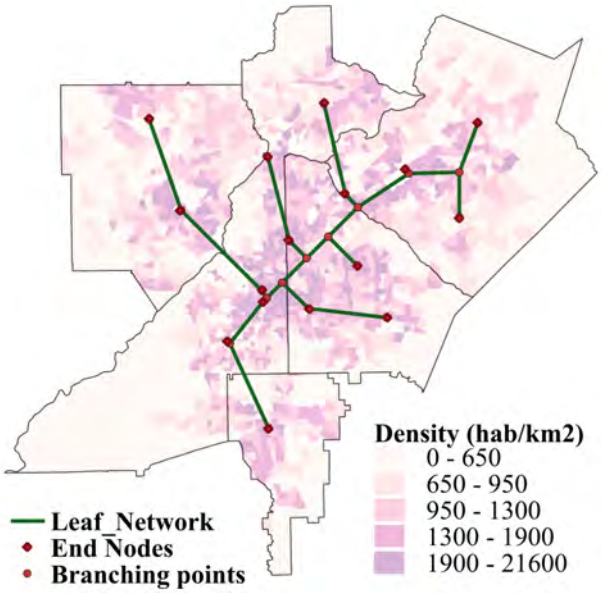

Figure 5.8: Leaf Venation (LV) network connecting the 15 discrete density of population centroids in Metro Atlanta.

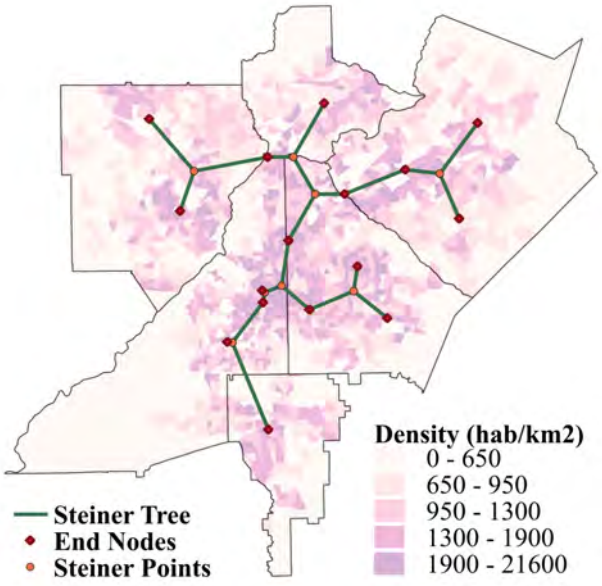

Figure 5.9: Steiner tree, resulting network connecting the 15 discrete density of population centroids in Metro Atlanta.

Pareto optimality

Following the same strategy as that described in subsection 5.4.2, we build the Pareto front that spans between the FN and the ST as shown in Figure 5.10.

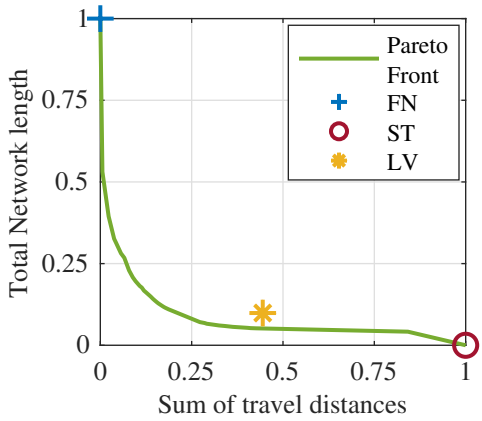

Figure 5.10: Pareto front and LV network for the transportation system built in the metropolitan area of Atlanta.

Results before normalization show that the FN and LV networks are respectively 106.7% and 10.53% longer than the Steiner tree. The sum of distances from the source is 14.86%

larger in the LV than in the FN and 25.10% larger in the ST than in the FN. The LV thus outperforms the ST for individual travel distances. The LV algorithm proves to be close to optimal, at a small distance from the Pareto front, with a high performance with respect to individual travel distances at the expense of network slightly longer than optimal.

Load balance and travel path efficiency

We measure the load balance following the method explained in subsubsection 5.4.3, except that the travel path between every pair of nodes is weighted according to the population associated to the nodes connected. Every edge that is a part of a travel path is assigned half of the total population represented by the two nodes that it connects. The total assigned population of an edge is the sum of the population assigned to that edge for all the paths that the edge is part of. The source node is assigned a weight of zero. Once the total congestion is calculated, the values are normalized by the total population. Figure 5.11 shows the load balance of the networks. The LV network shows a lower mean and interquartile range compared to the ST, suggesting that the edge load is slightly more homogeneous in the LV.

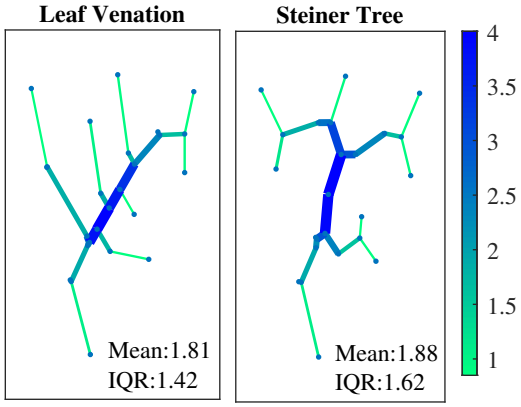

Figure 5.11: Load Balance as edge congestion. Edge thickness and color are proportional to edge congestion.

The travel path efficiency is evaluated as explained in subsubsection 5.4.3. Figure 5.12 shows the boxplots with the distribution of the path efficiency of the networks for both conditions. Additionally, Table 5.1 summarizes the mean and IQR values. The distribution of

the path efficiency from the source shows a small difference between the networks, with the LV exhibiting a better travel efficiency than the ST. On the other hand, the distributions of path ratio from all the centroids (attraction points and source) are similar for both networks, the LV performing slightly better than the ST.

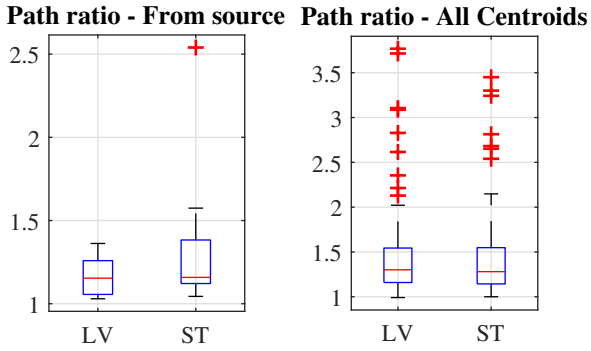

Figure 5.12: Path efficiency for the LV and ST networks in the metropolitan area of Atlanta.

Table 5.1: Path efficiency distribution, networks in metropolitan area of Atlanta.

Path Source	From Source		All Centroids	
	Mean	IQR	Mean	IQR
Leaf Venation	1.16	0.20	1.43	0.38
Steiner Tree	1.32	0.26	1.45	0.40

Population Served: Buffer method

We now assess the networks in terms of the population that they serve. We use the Geographic Information System (GIS) to calculate the total number of inhabitants (based on census tracts) that are within a distance of 2 km (walking distance) from the networks. The area bounded by the offset distance from the network is known as the buffer region. The buffer method is commonly used to evaluate transportation and service networks [235, 236].

Figure 5.13 presents the total population inside the buffer, the buffer area and the networks length. The three buffer variables exhibit similar trends: the FN reaches the largest

population, covers the largest area and has the highest network length. The LV is in second place and the ST is last. Interestingly, the values for the LV metrics are consistently about 10% higher than the ones of the ST. Figure 5.14 shows similar indexes, this time normalized by network length or buffer area. The population served per unit length shows that the most efficient networks are the ST (5,317 hab/km) followed closely by the LV (5,248 hab/km – 1.3% difference); the FN reached 3,687 hab/km – 30.7% difference. The area of the domain served by unit length of network is 4.03 km2/m for the ST, 3.99 km2/m for the LV (1% difference) and 2.82 km^2/m for the FN (30.1% difference).

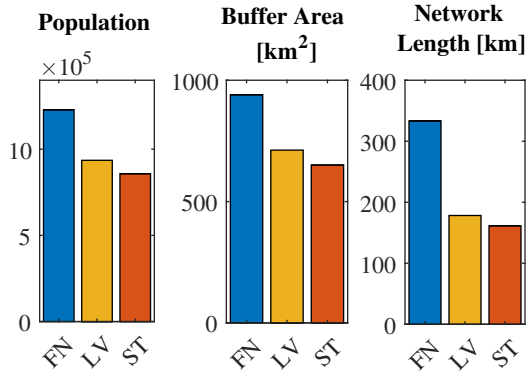

Figure 5.13: Indexes of population and area served by deployed networks. Metropolitan area of Atlanta.

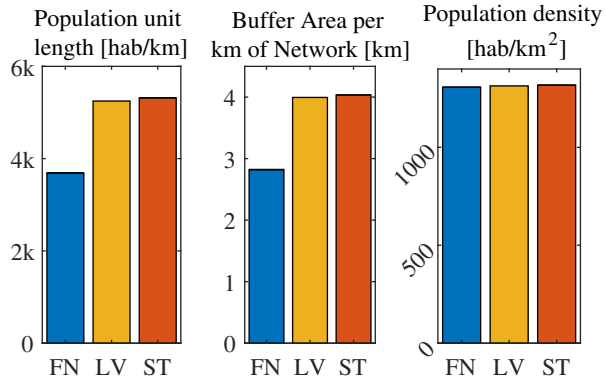

Figure 5.14: Normalized indexes of population and area served by deployed networks. Metropolitan area of Atlanta.

Lastly, the population density inside the buffer areas was very uniform along the networks, with an average of 1,313 hab/ km^2 and less than 1% difference among networks.

These results suggest that even though the resource concentration of the domain is the same for all the networks, the ST and the LV cover it more efficiently, both in terms of area covered and population served. The LV reached 10% more area and population than the ST, at the price of an increase in network length of 10%.

5.5.2 Gwinnett County

Gwinnett county is the second most populated county of the metropolitan area of Atlanta and is still not served by any railway to this date. We model the expansion of the MARTA railway network with the LV and the ST algorithms. By contrast with the modelling exercise presented for the whole Atlanta metropolitan area, the network expansion spans from a source node that may have a considerable effect on the overall urban network. Additionally, the edge capacity is not fixed or uniform.

We first discretize the population map of Gwinnett county by calculating the position of five weighted population centroids with a weighted k-means algorithm. The source node is represented by the current MARTA station that is the closest to Gwinnett county. The resulting LV and ST networks are shown in Figure 5.15 and Figure 5.16, respectively.

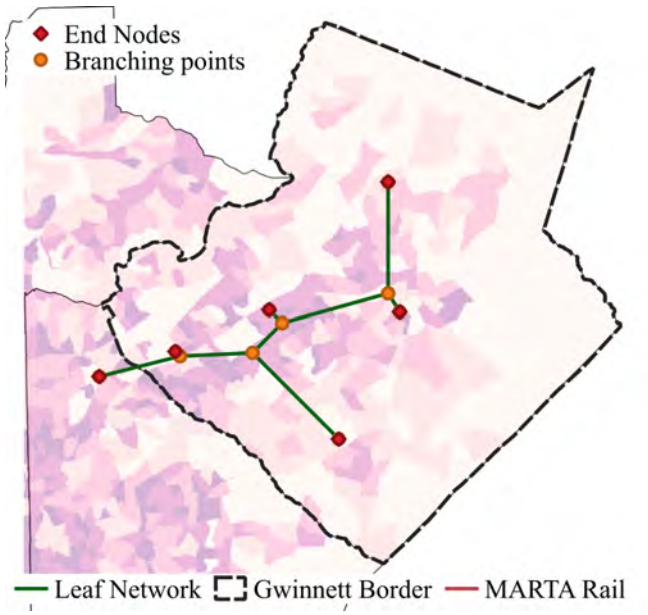

Figure 5.15: LV network deployed from the MARTA railway to five population centroids in Gwinnett county.

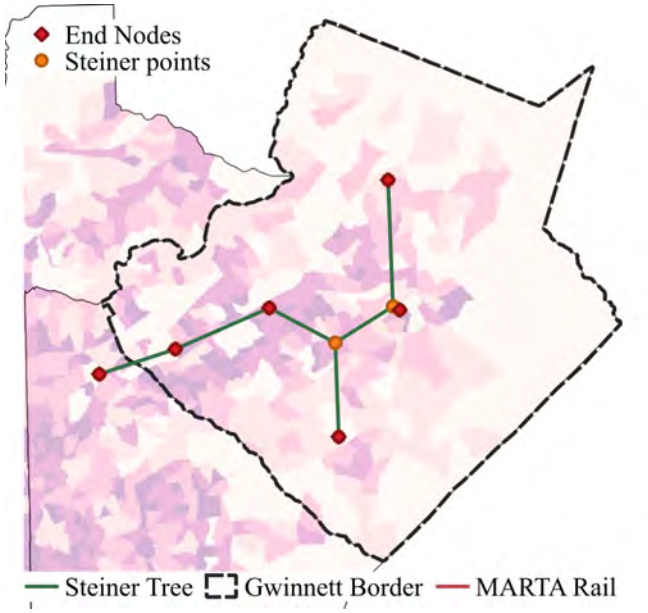

Figure 5.16: ST network deployed from the MARTA railway to five population centroids in Gwinnett County.

The total length of the LV is 6.1% higher than that of the ST (LV: 50.05 km; ST: 47.18 km). Regarding the population served, the LV reaches 10% more inhabitants than the ST (LV: 254,594 hab; ST: 231,496 hab). The area covered by the LV is 4.23% higher

than the ST. Table 5.2 summarizes the normalized indexes of population/area. We observe that even though both networks cover a similar area per unit length of network, the LV is more efficient, both in terms of population per unit length of network and in terms of total population reached. This is because the LV passes through regions with an increased density of population.

Table 5.2: Population and area reached by the LV and ST networks in Gwinnett County.

Network	Population per unit length [hab/km]	Area per unit length [km^2/m]	Population Density [hab/km^2]
ST	4906.1	4.22	1160
LV	5086.6	4.15	1225

By contrast with the cases simulated in previous sections, in which networks have edges of homogeneous capacity, we now consider that edge capacity is a design variable, and that, edge capacity should be proportional to its traffic load. We assume that the cost of each edge of the network is proportional to the product of its length by the load going through it (e.g. number of lanes multiplied by the length of the road). The total cost of the network is calculated as:

$$Cost = \sum_{i}^{n_E} C_i \cdot L_i \tag{5.3}$$

Where C_i is the traffic through the edge (i.e. the edge load, calculated as explained in Subsection 4.1.2), L_i the segment length and n_E is the number of network edges. We study the change in traffic distribution and network cost as a function of the weight (population) of the source node (W0). The influence of the source node on the network is proportional to its weight. We vary the weight of the source node from zero to a maximum value corresponding to the population of the four adjacent counties to Gwinnett. The population of the adjacent counties is 2,749,889, about 3.1 times the population of Gwinnett County (889,954). Figure 5.17 shows the traffic distribution for three different W0 values: zero,

Gwinnett's population, and the population of the adjacent counties. Networks are relatively independent from the rest of the railway system for homogeneous traffic loads, while for a concentrated load, the traffic is routed to a main vein in the neighborhood of the source node.

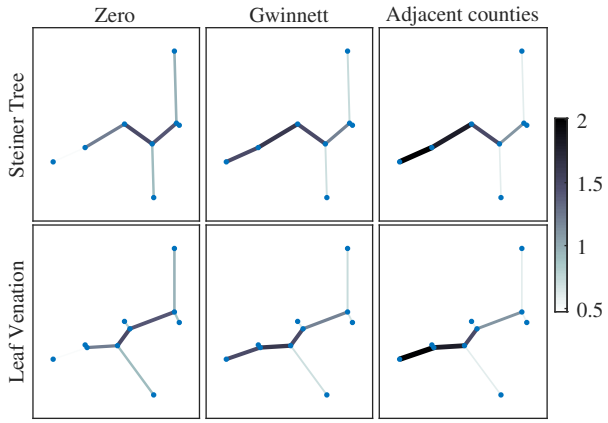

Figure 5.17: Traffic distribution as a function of source node weight for Gwinnett County. Edge thickness and intensity are proportional to their load.

Figure 5.18 shows the evolution of the total network cost and of the cost per unit length of network, as a function of the source node weight (W0) normalized by the population of the county. Results show that, for networks with low dependency on a source node, STs are more cost efficient than LV networks. Nevertheless, as networks become more dependent on a source node, LV networks outperform STs, yielding a lower total cost even though the total network length is always higher than the ST. In the current example, both networks have the same cost when the weight of the source point corresponds to 0.64 times the county's population. Lastly, it is important to mention that regardless of the source node weight, the normalized cost per unit length of network is always lower for LV networks, and the difference increases with the dependency from the source node.

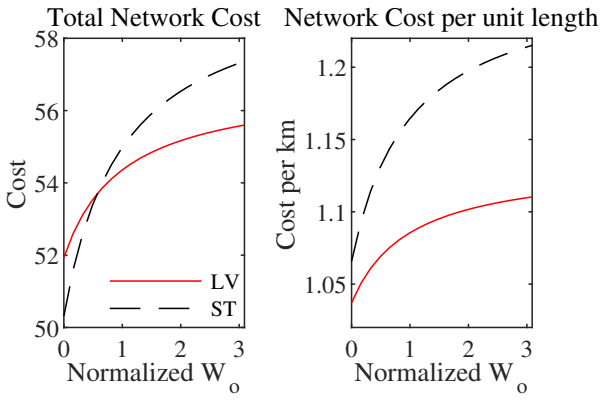

Figure 5.18: Total and unitary network cost as a function of source node weight - Gwinnett County.

5.6 Conclusions

Our simulation results show that the total length of LV networks is in average 10% larger than that of the minimum spanning tree (ST). Traffic distribution is slightly better in LVs than in STs, arguably because of LVs have a larger total network length. Additionally, LVs exhibit higher efficiency building paths from the petiole (source node) to auxin points (attraction points) than between pairs of auxin points, which is consistent with the natural transport function of a leaf. These improvements on traffic distribution and path efficiency are achieved at the expense of an increase in total network length. This evolutionary trade-off can be studied using Pareto optimality in future studies.

The length of LV networks of uniform edge traffic capacity spanning from the center of the city of Atlanta towards a set of 15 population centroids around its metropolitan area was 10% higher than the theoretical minimum, with travel distances in average 16% higher than the Euclidean distance for paths connecting the source node to the attraction points (vs 32% for the ST). Distances between attraction nodes were in average 43% higher than the Euclidean distance (vs 45% for the ST). LV networks thus outperform STs for transportation to and from a central node, while keeping the total network length and travel distances close to the optimal solution. Additionally, LV networks are as efficient as ST in terms of area and population reached per unit of network length.

A simplified problem of railway expansion was solved with the LV and ST algorithms, in which the capacity of the edges was proportional to their construction cost. Both networks reached a similar population per unit of network length (3.5% more population with the LV than the ST). The relative weights of the population centroids highly influenced the distribution of edges thickness of the network. The traffic load was higher for edges adjacent to the source node at which the railroad expansion was initiated. The network cost per unit of network length was always lower for LVs than STs. The cost difference increased with the weight of the central (source) node. This means that even though LVs exhibit a higher network length, their total cost is lower than that of ST for centralized networks. This is an interesting result for the development of Atlanta, GA and for the enhancement of the transportation networks in many other cities in the world, like New-York City (USA) or Paris (France).

We conclude that leaf venation algorithms can efficiently assist the design of engineered transportation networks. Nevertheless, our study was restricted to transport optimization and actual engineering design must consider other constraints including interference with current infrastructure, construction methods and operation limitations. We propose that leaf venation - inspired networks can be used to establish an initial design that can be refined based on environmental and engineering constraints. The advantage of using LV algorithms is that they achieve polynomial runtime instead of ST algorithms which are NP-complete. LVs are thus particularly suitable for determining initial network guesses that can then be iteratively optimized.

Conclusions of this study can be extended to multiple-connected networks in which secondary edges form loops. Network redundancy increases robustness at the expense of network cost. Additionally, LV algorithms can be used on continuous domains leveraging the assumption of discrete attraction points presented in this study. Lastly, there is an opportunity to expand LV algorithms to account for extra constraints; for instance, LV algorithms could be used to optimize routing or transport algorithms where the cost of the

network is a complex function considering land cost, edge capacity and network tortuosity, rather than just a function of the network length.

CHAPTER 6
EXPERIMENTAL STUDY OF THE DEFORMATION AND FAILURE MECHANISMS OF GRANULAR SOIL AROUND PRESSURIZED SHALLOW CAVITIES

6.1 Aim and Summary

The current chapter shows an experimental study on the deformation and failure mechanisms of cylindrical, shallow cavities embedded in a granular medium, which are relevant to many geotechnical applications, including tunneling and horizontal directional drilling. We study the effect of cavity length, vertical stress and soil density on such mechanisms using x-ray computed tomography.

Results showed that a closed failure surface developed around the cavities, beyond which, shear bands of elliptic paraboloid shape formed, extending from the bottom of the cavities all the way to the free surface. The plane strain assumption did not hold beyond the central portion of the longest cavity tested (L = 6D). The volumetric strain and porosity changes inside the shear bands showed significant dilation in dense specimens, but contraction in loose specimens.

The average orientation and the thickness of the shear bands were in agreement with those found in the literature for passive arching mechanisms (anchoring). Lastly, the orientation of the principal strains around the cavity followed a catenary shape, similar to that displayed in active trapdoor mechanisms. The content of this chapter has been submitted for publication in [42].

6.2 Introduction

Cavity expansion models are used in a wide range of applications in geomechanics, from in-situ testing [30, 31] to foundations and tunnel design [237, 238]. The pressuremeter test [32], the cone penetration test [33, 34] and the dilatometer test [35] all involve a pressurized cavity. In today's practice, extraction wells relevant to hydraulic fracturing and geothermal energy, deep tunnels, micro-tunnels and horizontal directional drilling projects [36] are designed from the cavity expansion theory. Analytical solutions typically assume plane strain, stress/strain coaxiality and absence of shear at the cavity wall. Stress distributions around holes were initially calculated within the theory of elasticity [239, 240]. Solutions were found for pressurized cavities by means of the superposition principle. Extension to plasticity [241] brings important mathematical challenges, especially in the absence of symmetries that reduce the dimensionality of the problem. An important tool used to find closed-form solutions to problems of cavities under anisotropic far-field stresses is complex mapping, as described by [242]. Analytical solutions differ by the choice of constitutive model and with it, drained or undrained conditions, and the combination of cavity shape and orientation of far field stresses (isotropic vs. anisotropic) [243].

Cavity expansion theories were developed for infinite media, and are only applicable when the size of the cavity is relatively small compared to the dimensions of the domain it is embedded in. Therefore, the effects of free surfaces are not accounted for, and the vertical and horizontal far field stresses can be considered uniform around the cavity, which typically results in isotropic or biaxial stress conditions. In a problem of shallow cavity expansion, these assumptions do not necessarily hold true, since the gradient of stresses around the cavity becomes significant and the geostatic stress field can no longer be modelled as a biaxial state. Furthermore, as observed in the results of the present study, the presence of a free surface close to the expanding cavity changes the failure mechanism of the soil, which transitions from a circular/elliptical plastic region around the cavity, to

a mechanism dominated by the development of shear bands starting from the cavity and reaching the free surface. Such failure mechanism resembles the well known passive trapdoor mechanism, first studied by Karl Terzaghi in 1936 [39] and later documented in a number of studies, e.g. [40, 41]. To date, few studies have treated the passive trapdoor mechanism triggered by the pressurization of a cavity. And more importantly, analytical solutions for shallow pressurized cavities developed within the theory of cavity expansion are inaccurate, because the failure mechanism is controlled by shear bands rather than by a plastic region enclosing the cavity. For example, current models of cavity expansion cannot predict the so-called blowout pressure in horizontal directional drilling [36, 37, 38].

In this chapter, we analyze the failure mechanisms around a shallow cylindrical pressurized cavity embedded in dry sand. A reduced-scale model was tested under biaxial stress conditions and imaged with an x-ray micro-x-ray scanner during pressurization, depressurization and re-pressurization to failure. Soil failure could not be explained by the cavity expansion theory and was instead dominated by passive anchoring. We analyzed the influence of vertical stress, soil density and cavity length on cavity deformation, soil porosity changes and displacement/strain fields around the cavity. We first describe our experimental protocols along with the materials tested. Second, we present our image analyzes and we interpret the cavity deformation and the soil strain field. Third, we discuss the failure mechanisms and we evaluate the plane strain assumption. Lastly, we summarise our conclusions.

6.3 Materials and Methods

6.3.1 Experimental Setup

The experimental setup is shown in Figure 6.1. A cylindrical container held the soil specimen and applied the surcharge vertical stress (σ_v). The soil specimen was 10 cm in height and diameter. The inflatable probe that was used to model the pressurized cavity was placed

horizontally inside the soil. A syringe pump was connected to the probe for pressurization. All the components were made of plastic, since metals cause significant attenuation of the x-rays.

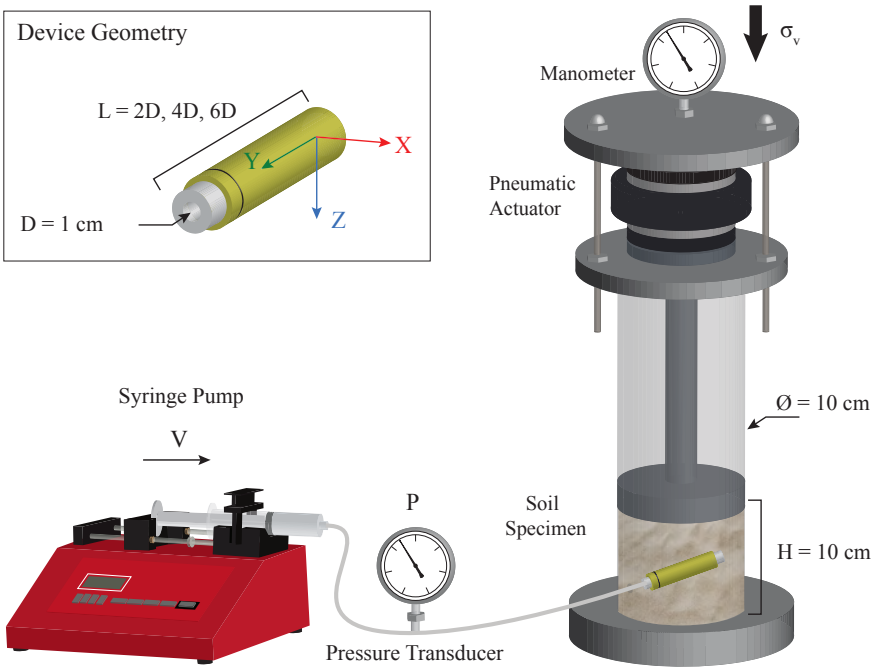

Figure 6.1: Experimental Setup. Soil specimen placed inside scanning region. Pressure and volume monitored during expansion steps and scanning.

We use a Cartesian coordinate system in which the Z-axis (in the vertical direction) is oriented along the axis of the cylindrical container. The datum is placed at the top of the soil specimen and the Z axis points downwards, i.e. Z increased with depth. The Y-axis is oriented along the axis of the (initially) cylindrical probe/cavity, with the origin at the end of the probe furthest away from the pump. The X-axis is normal to the Y and Z axes, as shown in Figure 6.1. From this point onward, we will refer to three cross-sections, all passing through the mid-point of the device along the Y axis. The cross-section along the XZ plane, in which the device appears as a circle/ellipse, will be referred as the front view. The cross-section along the YZ plane will be referred as the side view, and the section along the XY plane is referred as the top view.

The soil container is made of a plexiglass cylinder with an internal diameter of 10 cm,

a wall thickness of 1 cm and a height of 40 cm. The loading system fastened at the top of the container consists of a pneumatic actuator connected to a plastic piston that transferred the load to the top of the soil specimen. The pressure applied by the actuator was regulated using a manometer connected to a line of pressurized air. The use of a relatively long plastic piston to transfer the load was necessary, to guarantee that none of the metallic parts of the pneumatic actuator fell within the region of the set up being scanned. Three different vertical pressures were applied in the experimental campaign: $\sigma_{V1} = 3.5kPa$, $\sigma_{V2} = 6kPa$ and $\sigma_{V3} = 8kPa$.

We used HN31 sand [244], a fine, angular and uniformly graded sand with particle sizes between $0.16mm$ and $0.63mm$, and a mean particle size (d_{50}) of $0.34mm$, and a coefficient of uniformity of 1.70. The specific gravity (G_s) of the material was 2.65. The minimum and maximum unit weights of the material are 13.24 and $15.99kN/m^3$ [245, 244], which correspond to void ratios of 1.00 and 0.66 or porosities of 50.0% and 33.7% respectively, and will be used as the bounds to define the relative density (D_R) measurements. The specimens were dry pluviated in two layers. The first layer extended from the bottom to a height of $5cm$, where the probe was placed horizontally at the center of the cylinder. The second layer was formed to reach a total height of $10cm$. Specimens were pluviated from different elevations to achieve two different relative densities D_R of 22% ("loose specimens") and 65% ("dense specimens").

The probes were $10mm$ in diameter and were similar to a single chamber Menard pressuremeter. Holes were drilled with regular spacing in an acrylic cylinder of $9.52mm$ in diameter that was wrapped within a thin latex membrane with a thickness of approximately $0.25mm$. The membrane was clamped at the ends of the cylinder with o-rings and custom made gaskets. One of the ends of the probe was connected with a tube fitting to a relatively rigid polyurethane tubing connected to the inflation setup. The other end was closed tightly using a plastic screw to avoid leaks. We tested probes of three different lengths: 20, 40 and $60mm$, corresponding to 2, 4 and 6 times the probe diameter.

The inflation setup consisted of a syringe pump set at a constant rate of $0.2cm^3/min$. Probes were inflated with water, which can be assumed incompressible at the range of pressures tested (below $200kPa$). A high-grade syringe pump was connected to the device using polyurethane tubing. The length of the tubing was kept to a minimum in order to minimize volume and pressure loss. A pressure transducer (Omega PX26) with a capacity of $200kPa$ was connected to the pressure line, using a T-shape, to monitor the internal pressure of the system, as shown in Figure 6.1.

6.3.2 Testing procedure

Before every test, each probe was purged to make sure that no air bubbles were trapped inside. Then, following previous recommendations [30], two different calibration steps were performed before each test. First, the probe was hung vertically and inflated up to three times its initial volume. The obtained pressure vs. volume response of the probe corresponds to the resistance of the membrane and the tubing system, which we later deducted from the response of the expansion inside the soil specimen. Second, we measured the compliance of the system, e.g. the deformation of the tubing assembly. The probes were placed inside a tight fitting metal tube (considered rigid at the range of pressures tested) and then pressurized until a maximum pressure of 200 kPa was reached. Figure 6.2 shows typical calibration curves. Incidentally, measuring the compliance of the system helped identify leaks and trapped air bubbles.

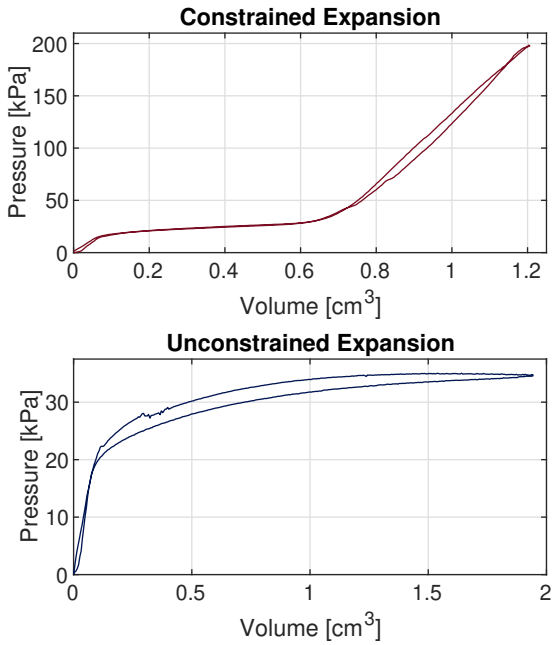

Figure 6.2: Typical probe calibration curves. "Dense" sand, probe with L = 2D and vertical stress $\sigma_{V1} = 3.5kPa$

After each specimen was prepared, the container was placed in the x-ray scanner and scanned three times. The first baseline step scanned the specimen before the inflation of the probe. Next, the probe was inflated until the pressure reached a peak or stabilized, and the inflation stopped. A second scan was taken then. Lastly, the device was further inflated until its volume reached three times its initial volume, and then a third scan was performed. We conducted a total of 13 tests, 12 of them combining two relative densities (D_R 40% and 75%), three probe lengths (2, 4 and 6 times the probe diameter) and two surcharge pressures ($\sigma_{V1} = 3.5kPa$ and $\sigma_{V2} = 6kPa$). A last test was conducted on a specimen with D_R 40%, probe length $6D$ and vertical surcharge $\sigma_{V3} = 8kPa$.

6.3.3 Image acquisition

The scanning volume was a prism with a height of $90mm$ and a cross section of $119mm$ by $119mm$. The resolution of the acquired images was controlled by the voxel size, which we set as $70\mu m$ (approximately one fifth of d_{50}). After cropping the acquired images, we obtained stacks of 1430-by-1430-by-1000 voxels, containing the upper $70mm$ of the soil

specimen. The present section includes the strictly necessary details about the process of image acquisition using x-ray computed tomography (CT-scanning). For further details and a comprehensive description of the Digital volume correlation (DVC) procedures, we refer the readers to [246].

6.3.4 Image Analysis: Cavity Deformation

Images were segmented to analyze the deformation of the probe (or cavity). The grey value that best separated the probe from the soil grains was used as a threshold to binarize the images and exclude the probe domain from the soil deformation analysis. The binarization was performed sequentially, starting from the horizontal slice that passes through the center of the probe (of known location). We identified the horizontal cross section of the probe as the largest connected component after binarization. Then, the raw cross section was smoothed by eroding and dilating its contour, with a structuring element of 10 voxels in radius (two times the mean grain size). Then, two independent loops were run to sequentially add horizontal cross sections to the segmented 3D cavity volume towards the top and bottom of the specimen. In order to ensure continuity of the segmented volume, the connected component of a subsequent slice must share at least one voxel location with the previous slice. The cross sections of the cavity reduced in size as the distance from the middle of the device increased, until the cavity footprint vanished, at which point, the segmentation was completed. Then, the centroid, area, major and minor axes of the convex (ellipsoidal) cross-sections of the cavity were stored for further analysis.

6.3.5 Image Analysis: Soil Deformation

Before analysing soil deformation, corners of the cylindrical container were masked together with the segmented cavity volume. Soil displacement fields were computed using the open-source, python package SPAM [247]. The local DVC algorithm splits the 3D image into sub-volumes in the reference configuration, and for each subvolume iteratively

solves for the linear transformation function that offers the best match with the greyscale texture in the deformed configuration. The result is thus a transformation function for each subvolume which includes a 3D displacement vector whose accuracy is well below the pixel. Sub-volumes are centered around a given node, and are spaced along the three dimensions of the image. The size of the sub-volumes is controlled by the window size (WS). We set the WS as 33 voxels (about 6.4 times the average grain width - d_{50}), while the distance between nodes (known as node spacing) was 24 voxels (4.8 times d_{50}) in all three directions, meaning that there is some overlapping of subvolumes.

The obtained displacement fields had convergence rates above 98%, i.e. less than 2% of the nodes could not be correlated between consecutive images. The strain fields are computed using Q8 shape functions to compute the displacement gradient tensor (**F**), which is a measure of the local displacement variations. We used a finite strain formulation given the significant distortion of the soil matrix. Local measurements of porosity were performed on every image. The resolution of our images did not allow a binary segmentation between grain and void phases and therefore a linear greyscale calibration (non-binary) classification was used. The calibration between grey values and porosity was performed for each test (essentially the identification of the grains and void greyvalues), allowing porosity to be computed on appropriately defined subvolumes. Similar to the displacement and strain fields calculation, the voxels outside of the container and inside of the cavity were excluded from the analysis.

6.4 Results

6.4.1 Cavity response

Pressure vs. Volume response

The pressure - volume response of the cavities was monitored during the tests and was corrected to account for the compliance and resistance of the membrane around the probe. Figure 6.3 shows the resulting response of the cavities for the different probe lengths, soil densities and surcharge vertical stresses tested. The observed loss of pressure at constant volume (vertical lines in the cavity response) corresponds to the moment at which the inflation step was stopped to acquire the first tomography. The drops in pressure corresponds to the scanning periods, which lasted about an hour.

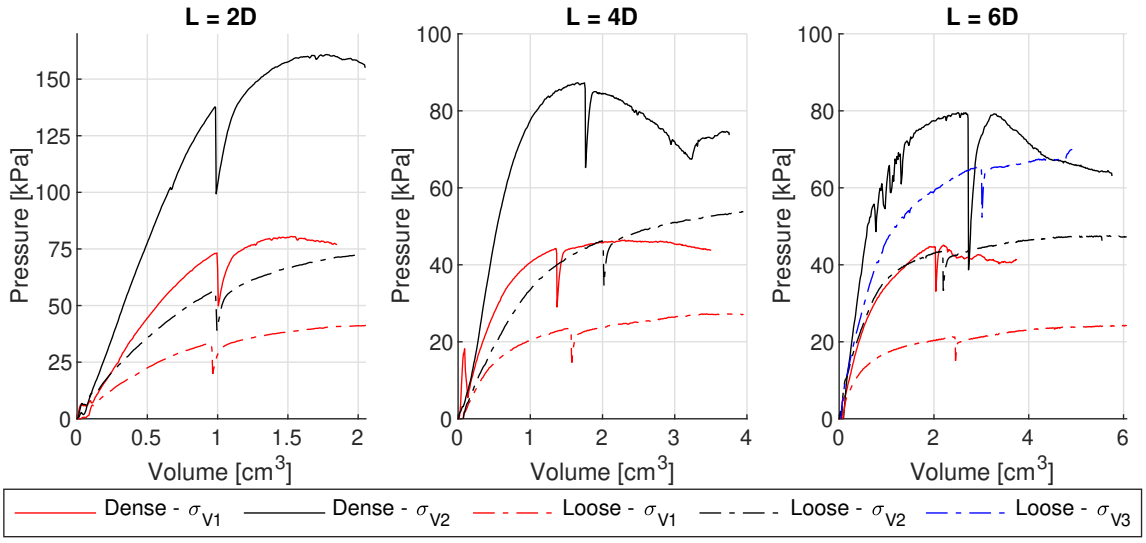

Figure 6.3: Pressure vs. Volume response of the cavities. Dashed and solid lines correspond to experiments in loose and dense soil respectively. Different surcharge vertical stresses are shown in different line hues. Note that the pressure scale in the left-most plot is different to the other two, due to large differences in magnitude range.

The two longer cavities (L = 4D, 6D) exhibited similar pressure-volume responses, while the shorter cavities showed significantly higher pressures as a result of the increase

in volume. Dense soils exhibited a stiffer response that reaches a peak resistance, and decreases afterwards, typical of dilative materials. Conversely, loose specimens showed a more ductile response, with an asymptotic behavior, typical of contractive specimens. In the same way, a higher vertical surcharge pressure resulted in higher maximum resistance of the specimen.

We fitted a multivariate linear regression with the variables (relative density, vertical surcharge and device length) as predictors, and the peak cavity resistance (maximum pressure value recorded during the test) as the response. All the variables were standardized (mean of 0 and standard deviation of 1). The linear regression was obtained as follows:

$$P = 0.66\sigma_v + 0.58D_R - 0.49L \qquad R^2 = 0.82 \tag{6.1}$$

The regression in Eq Equation 6.1 shows that 82% of the variability in the soil resistance can be explained from the sum of the individual contributions of the tested variables. The largest influence in the soil resistance (P) comes from the surcharge vertical stress (σ_v), followed by the relative density of the soil (D_R), both of which are positively correlated to P. Conversely, the device length (L) is negatively correlated to the peak response, meaning that shorter devices yield a higher peak resistance.

Cavity deformation

The anisotropic stress conditions around the cavity cause a non-homogeneous deformation of the cavity during its expansion. To characterize the shape of the cavity, which becomes elliptical during expansion, we calculated the eccentricity (e) of the cross-section, calculated as $e = 100(M/m - 1)$ where M and m correspond to the major and minor axis of the cavity cross section, respectively. Figure 6.4 shows the differences in cross-section area and shape along the device axis for each cavity length, soil density and surcharge stress tested.

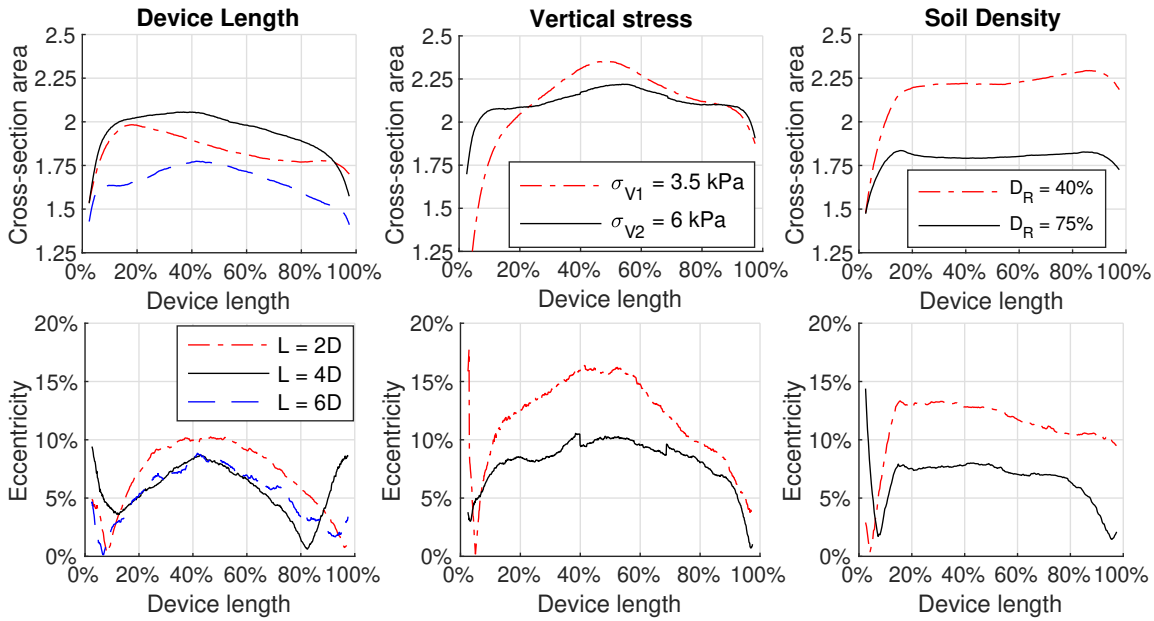

Figure 6.4: Effect of device length, surcharge vertical stress and soil specimen density on the cross-sectional area and shape of the pressurized cavities. The cross sectional area is normalised by the initial cavity area. Eccentricity is calculated as the ratio between the major and minor axis of the cavity cross-section minus one. The left column (Device length specimens) shows results of experiments performed on loose specimens under a vertical stress σ_{V2}. The middle column (vertical stress specimens) corresponds to a device length of 6D in loose soil. The right column (Density specimens) shows results for devices with a length equal to 2D and a vertical stress of σ_{V2}.

Results suggest that density has the highest influence on the area of the device, showing that larger cavity deformation and higher eccentricity occur in looser specimens. As expected, the cross sections towards the middle of cavity, furthest away from the circular clamps, exhibit the largest eccentricity. Similarly, results suggest that a lower vertical stress promotes eccentricity. Table 6.1 shows the mean cross-sectional area for each test, and the ratio between the maximum and mean areas, as a measure of homogeneity along the device. Similarly, Table 6.2 shows the mean eccentricity and the ratio between the maximum and mean eccentricity values along the cavity.

Table 6.1: Cavity cross-sectional area. The M index corresponds to the mean cross sectional area along the device length, normalised by the initial cross-section of the cavity. The R index corresponds to the ratio between the maximum and the mean cross-sectional areas along the cavity as a measure of homogeneity.

Stress		σ_{V1}			σ_{V2}		
Length		**2D**	**4D**	**6D**	**2D**	**4D**	**6D**
Loose	**M**	2.33	1.70	2.10	2.18	2.21	2.11
	R	5.54	6.81	11.72	4.75	3.05	4.87
Dense	**M**	1.84	1.95	1.66	1.79	1.94	2.11
	R	7.65	5.60	6.88	2.43	3.90	6.95

Table 6.2: Cavity cross-sectional shape. The M index corresponds to the mean cross sectional eccentricity in percentage. The R index corresponds to the ratio between the maximum and the mean cross-sectional eccentricity along the cavity as a measure of homogeneity. Eccentricity is calculated as the ratio between the major and minor axes of the cavity, minus the unit, and shown as a percentage.

Stress		σ_{V1}			σ_{V2}		
Length		**2D**	**4D**	**6D**	**2D**	**4D**	**6D**
Loose	**M**	13.1	15.3	11.9	11.2	12.9	8.2
	R	30.1	10.2	38.0	19.2	34.0	29.3
Dense	**M**	7.00	5.74	5.31	6.59	8.35	9.80
	R	46.5	51.8	66.9	21.5	44.4	59.2

Results show that even though the mean cross sectional area is relatively similar among tests, higher soil density and vertical stresses restrict the expansion of the cavity, resulting in a more homogeneous expansion along the cavity length. Denser soils reduce the overall eccentricity of the device, but the intermediate portion of the cavity still remains more eccentric than its neighboring regions.

6.4.2 Soil Strain field

Soil strain is quantified by means of the right-hand stretch tensor (**U**), obtained after a polar decomposition of the gradient transformation tensor (**F = RU**). Regardless of the test variables, the strain fields exhibit two overlapping mechanisms: a zone of high strain around the cavity itself, reminiscent of cavity expansion; and a failure surface that extends from the bottom of the cavity to the free surface, typical of a passive trapdoor mechanism. An additional disturbance zone above the cavity, inside the region enclosed by the failure surface, is also observed in later stages of the expansion. Figure 6.5 exemplifies these mechanisms displaying the strain magnitude (2-norm of the stretch tensor **U**).

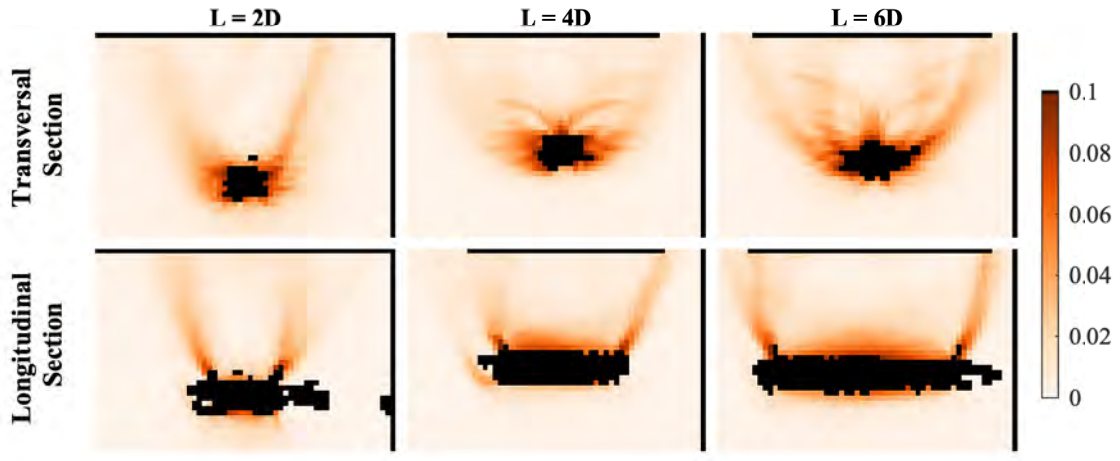

Figure 6.5: Strain magnitude calculated as the Euclidean norm of the stretch tensor (**U**). Values shown for tests in "dense" specimens and vertical surcharge σ_{V1}.

Deviatoric Strain

The deviatoric strain scalar is calculated as the Euclidean norm of the deviatoric part of the stretch tensor (**U**). Results in the loose specimens suggest that the pressurization of shorter cavities (L = 2D), closer to a spherical shape, produces a concentric region of strain with less prevalent shear bands, contrary to longer cavities that yield shear bands. Conversely, dense specimens exhibit a higher concentration of strain along the shear bands when compared to looser specimens, even in the shorter cavities. Figure 6.6 and Figure 6.7 show the

deviatoric strain fields in loose and dense specimens respectively.

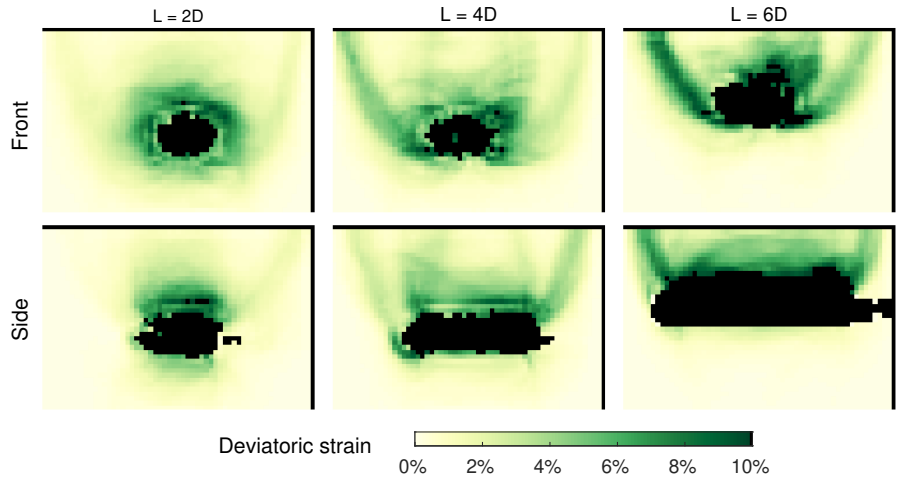

Figure 6.6: Euclidean norm of the deviatoric part of the stretch tensor in loose specimens. Shown fields correspond to the tests with a vertical surcharge σ_{V1}.

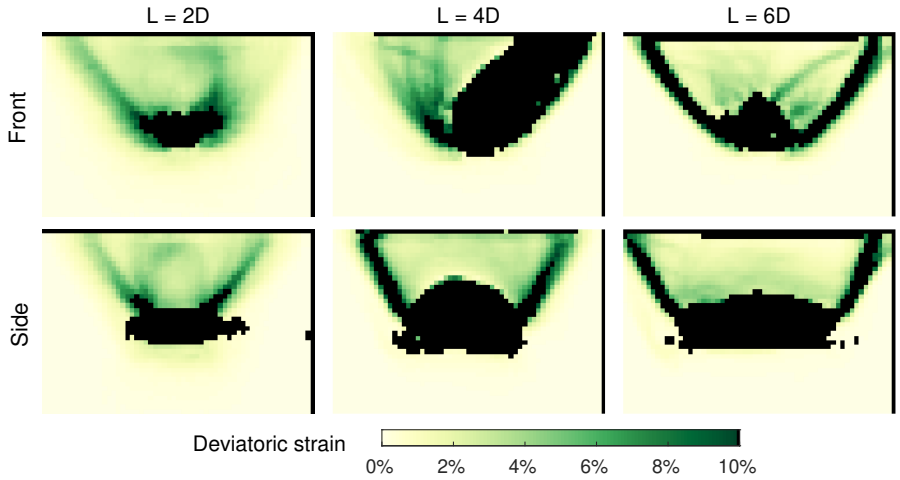

Figure 6.7: Euclidean norm of the deviatoric part of the stretch tensor in dense specimens. Shown fields correspond to the tests with a vertical surcharge σ_{V1}.

Volumetric Strain

The volumetric strain is calculated as $J - 1$, where J is the Jacobian of the stretch tensor, with dilation counted positive and contraction counted negative. In loose specimens, a zone of dilation generates around the cavity, and contraction is observed along the shear bands, even before they have fully developed. Conversely, in dense specimens, soil dilation is

predominant both around the cavity and along the shear bands. It is also worth noting that the magnitude of the volumetric strain is significantly higher in the dense specimens than in the loose specimens. Figure 6.8 and Figure 6.9 exemplify the distributions of volumetric strain in loose and dense specimens respectively.

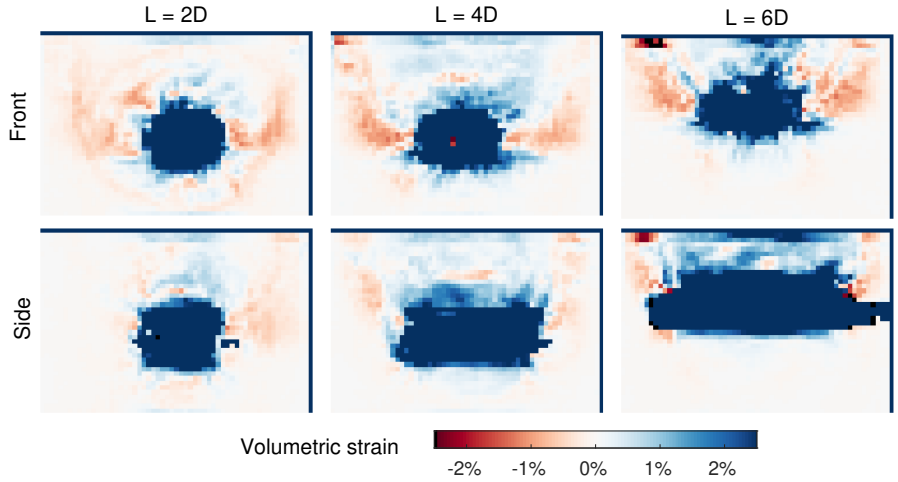

Figure 6.8: Volumetric strain in loose specimens. The volumetric strain is calculated as $J-1$, where J is the Jacobian of the stretch tensor. Negative and positive strains correspond to contraction and dilation, respectively. Shown fields correspond to the tests with a vertical surcharge σ_{V1}.

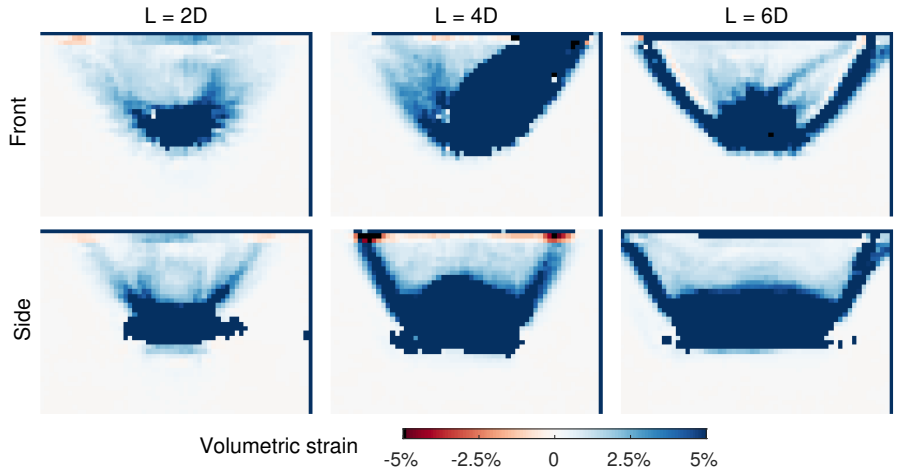

Figure 6.9: Volumetric strain in dense specimens. The volumetric strain is calculated as $J-1$, where J is the Jacobian of the stretch tensor. Negative and positive strains correspond to contraction and dilation, respectively. Shown fields correspond to the tests with a vertical surcharge σ_{V1}.

Porosity change

Calculation of porosity change confirms that the material along the shear bands compacts in loose specimens and dilates in dense specimens. A dilation region around the pressurized cavity was present in every test, irrespective of the specimen density, applied stress or cavity length. Figure 6.10 and Figure 6.11 show the change in porosity for loose and dense specimens, respectively.

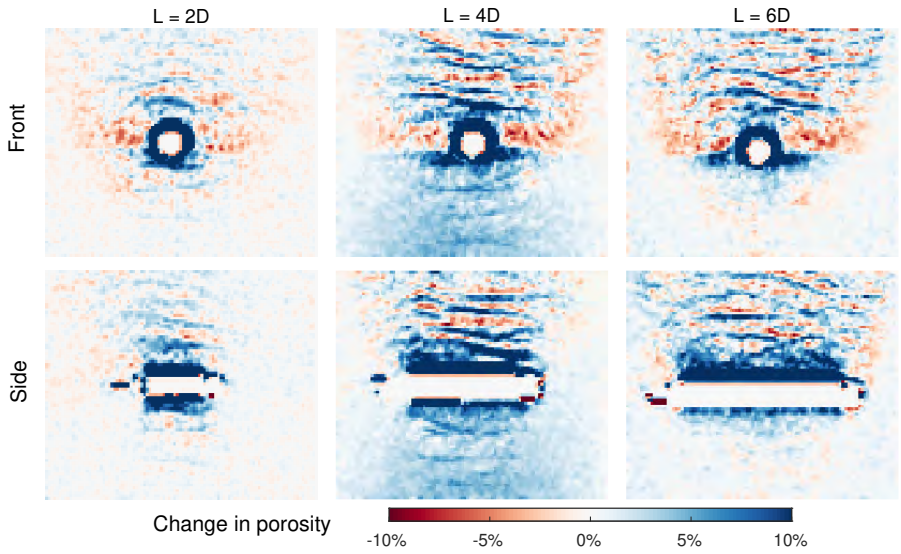

Figure 6.10: Change in porosity in loose specimens, calculated as the absolute porosity change between the final expansion step and the initial state. A negative change in porosity corresponds to densification and a positive change, to loosening. Shown fields correspond to the tests with a vertical surcharge σ_{V1}.

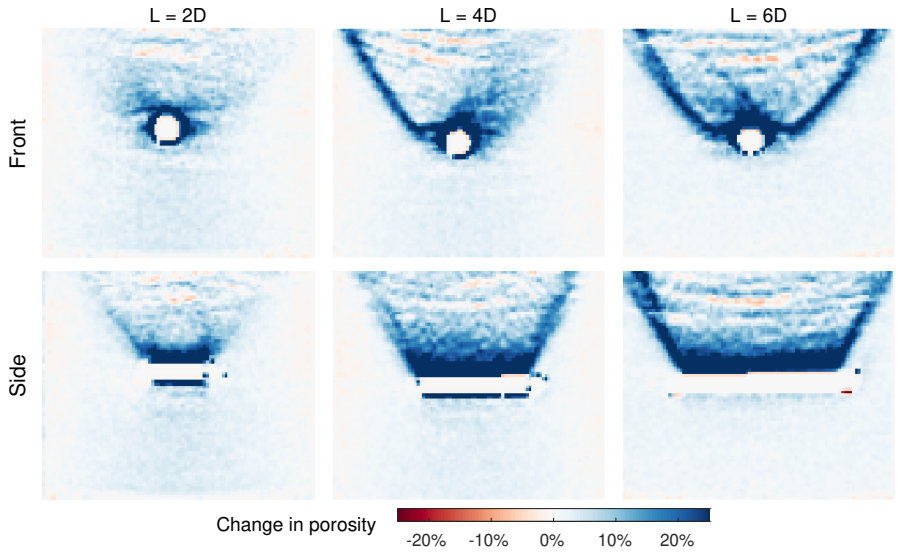

Figure 6.11: Change in porosity in dense specimens, calculated as the absolute porosity change between the final expansion step and the initial state. A negative change in porosity corresponds to densification and a positive change, to loosening. Shown fields correspond to the tests with a vertical surcharge σ_{V1}.

6.5 Discussion

6.5.1 Failure surface

The shear bands that develop around the cavities exhibit a concave curvature, enclosing a failure zone that extends from the cavity to the surface. In the following, we characterize the shape of the failure surface in the XZ plane (front view), study the thickness of such shear bands and quantify their orientation using a straight line approximation.

Failure zone shape characterization

The shear zone generated during the expansion describes a bell shape, with its vertex below the invert of the cavity and convex branches extending to the surface (see Figure 6.5). In order to characterize such shape, we manually picked the boundary of the shear region from the front cross-section of the maximum shear strain (absolute difference between the major and minor principal strain increments) using a threshold of 7.5%. The obtained contours

were then fitted to catenary curves, of equation:

$$z = a * \cosh\left(\frac{x - x_c}{a}\right) - z_c \tag{6.2}$$

where x_c and z_c correspond to centering coordinates in the x and z axis respectively, and a is a shape parameter. Table 6.3 shows the obtained values for the shape parameter a. The coefficients of determination were $R^2 >= 0.9$.

Table 6.3: Shape characterization of the failure zone: values of the shape parameter of the catenary a for the different experimental variables. Smaller values of a correspond to curves with a higher curvature.

Stress	σ_{V1}			σ_{V2}			σ_{V3}
Length	**2D**	**4D**	**6D**	**2D**	**4D**	**6D**	**6D**
Loose	22.0	27.4	29.2	27.7	28.1	30.2	34.1
Dense	22.7	25.4	29.0	27.7	25.3	32.4	-

The shape of shear zones has been studied for similar loading scenarios, including uplift capacity of anchors, uplift resistance of buried pipes, and limit analysis of shallow tunnels. For instance, [248] studied the pullout of flat, circular anchors and described the shape of the failure surface as a "spatial funnel", with radially symmetric convex boundaries. Similarly, [17] obtained similar results for the failure surface of root-like anchors using CT-scanning. Literature on the uplift resistance of buried pipes, in which circular inclusions are pulled vertically assuming plane strain conditions, have shown concave shapes starting between the waist and the crown of the pipes [249]. More recently, [250] developed a limit analysis formulation for pressurized shallow tunnels, which resembles the loading conditions in the present study, but still shows failure planes that do not completely enclose the cavity.

Contrary to the results reported in the literature ([249, 250]), our experimental results show that the shear region is convex and completely encloses the cavity. The results from Table 6.3 suggest that higher vertical stress results in less concave failure zones (which tend

to be more straight lines). Shorter cavities in loose soil exhibit a more ductile response, with no evidence of peak resistance or shear bands and result in highly curved failure zones, which agrees with the observations by [249]. The stark difference of failure mechanism between short (L=2D) and longer cavities (L=4D and L=6D) coincides with the difference in the pressure-volume response (Figure 6.3).

Orientation and thickness of shear bands (side view)

For each test, shear bands were identified manually from the maximum shear strain field measured in the front cross section. We selected a threshold of 7.5% shear strain, which allowed us to clearly identify localized shear bands. We computed the medial axis of each shear band. The medial axis, also known as skeleton, is a common feature used in image processing, and corresponds to the set of points inside the region which are equidistant to more than one edge of the region. The distance from the medial axis of the shear band to its boundaries corresponds to half the shear band width. This distance, initially calculated in pixels, was transformed into millimeters and compared to the mean grain size of the soil ($d_{50} = 0.35mm$).

Results, provided in Table 6.4, show that the thickness of the shear bands is between 8.1 and 20.6 times the mean grain diameter, without clear influence of the tested variables (soil density, surcharge stress and probe length) in the values. These findings agree with with other experimental results reported in the literature, in which the shear band thickness is between 5 and 20 times the mean grain size regardless of test variables [251, 252, 245].

Table 6.4: Shear Band thickness (width), measured relative to the mean grain size diameter (d_{50}).

Stress	σ_{V1}			σ_{V2}			σ_{V3}
Length	**2D**	**4D**	**6D**	**2D**	**4D**	**6D**	**6D**
Loose	-	12.8	18.1	-	17.6	18.1	16.2
Dense	8.1	20.6	13.8	11.4	-	17.1	-

Next, the orientation of the shear bands was calculated. Since the obtained shear bands are curved, specially in the neighborhood of the cavity, we used the two thirds of the medial axis closest to the free surface to calculate the orientation angle of the shear bands. That portion of the medial axis was then fitted using a linear regression and the slope was transformed into an angle measured from the vertical (Z direction). Obtained results are shown in Table 6.5.

Table 6.5: Shear band orientation. Values in degrees measured from the vertical. Angles were obtained from the linear fit of the medial axis of the shear band, discarding the highly curved section adjacent to the cavity.

Stress	σ_{V1}			σ_{V2}			σ_{V3}
Length	**2D**	**4D**	**6D**	**2D**	**4D**	**6D**	**6D**
Loose	-	24.8	28.5	-	26.8	34.4	40.9
Dense	31.1	29.9	38.2	39.5	-	38.9	-

Note that the missing values in Table 6.4 and Table 6.5 correspond to the tests with the shortest cavity (L = 2D) in loose soil, which did not exhibit well defined shear bands, while the test with variables (L = 4D, σ_{V2} and dense soil) showed significant distortion and therefore it was not possible to identify the shear bands.

According to previous studies on the uplift resistance of buried pipes ([253, 249, 254]), the inclination angle of the shear bands (measured from the vertical) is equal to the dilation

angle (ψ), which depends on the stress state and density of the soil mass. Results reported in Table 6.5 are in agreement, showing that higher vertical surcharge stress and increased soil density result in higher angles of orientation of the shear bands.

Strain and porosity change inside the shear bands

We calculated the volumetric strain and the change in porosity inside the shear bands. Average values are reported in Table 6.6.

Table 6.6: Volumetric strain and change of porosity inside shear bands. V - Mean volumetric strain inside the shear bands [%] P - Mean porosity change inside shear bands. [%]

Stress		σ_{V1}			σ_{V2}			σ_{V3}
Length		**2D**	**4D**	**6D**	**2D**	**4D**	**6D**	**6D**
Loose	V	-	-0.07	0.52	-	-0.51	0.02	0.77
	P	-	-0.33	-1.63	-	-1.11	0.92	0.28
Dense	V	3.47	3.41	3.22	2.25	-	5.36	-
	P	8.79	9.61	7.68	5.43	-	14.75	-

Results from Table 6.6 show a linear relationship ($R^2 = 0.96$) between the change in porosity (and therefore change in void ratio) and the volumetric strain generated by the expansion of the cavity inside the shear bands. Additionally, it can be observed that the changes inside the shear bands in loose specimens are smaller in magnitude compared to dense specimens. The soil along the shear bands in loose state slightly dilates or contracts, while it significantly dilates when initially in dense state.

Table 6.7 shows the final void ratio achieved in the shear bands. Results show that the terminal void ratio (e) inside the shear bands reaches values above the value ($e = 1$) reported for the minimum density [245, 244], which is in agreement with the findings of [255], who showed that the density of the material along the shear bands can reach values below the minimum used to calculate relative density. Similarly, [256] found increments

of up to 24.7% in the void ratio inside shear bands in triaxial tests, while we found average increments of 22.5% 68.8% in loose and dense specimens respectively.

Table 6.7: Final void ratio inside shear bands [-]. Average values inside identified regions corresponding to shear bands in the different tests. Initial void ratios (e) for the loose and dense specimens were 0.91 and 0.78 respectively.

Stress	σ_{V1}			σ_{V2}			σ_{V3}
Length	**2D**	**4D**	**6D**	**2D**	**4D**	**6D**	**6D**
Loose	-	1.19	1.20	-	1.10	1.03	1.06
Dense	0.89	1.52	1.18	1.31	-	1.68	-

Figure 6.12 shows the change of the void ratio along the shear bands over the different expansion steps. Several authors have proposed that the void ratio inside shear bands converges to a critical value [257, 258] and [259, 260] further proposed that such value depends on the initial void ratio of the soil. Results from Figure 6.12 further demonstrates the difference of behavior between loose and dense specimens: loose specimens exhibit signs of convergence, while dense specimens do not. In addition, void ratios inside shear bands are significantly higher in dense specimens compared to loose ones and are inversely proportional to the vertical surcharge stress in our experiments.

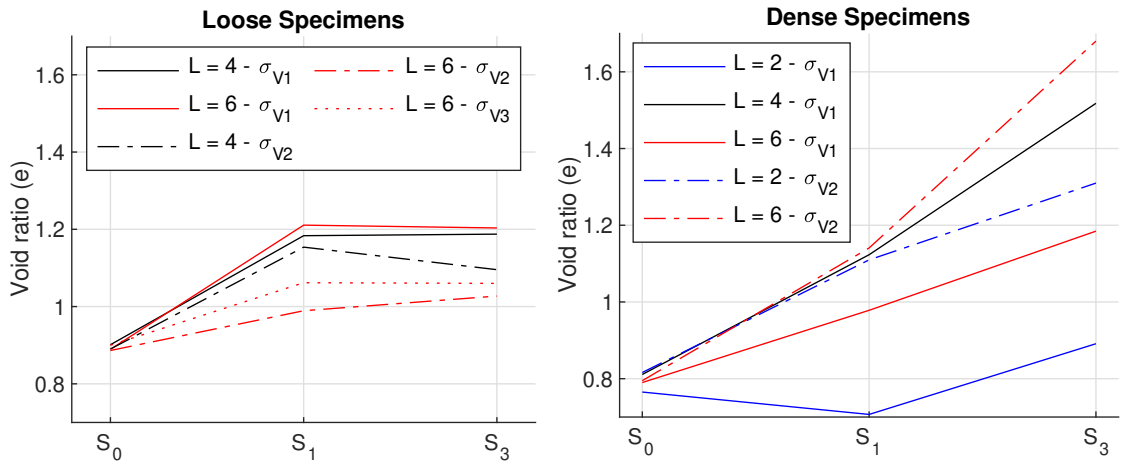

Figure 6.12: Change in void ratio inside shear bands during the expansion steps. Average void ratio along the location of the shear bands before expansion (S_0), and for the subsequent expansion steps S_1 and S_2.

6.5.2 Orientation of principal strain increments

Principal strain increments and their respective orientations are respectively obtained as the eigenvalues and the eigenvectors of the stretch tensor. The obtained strain orientations were similar for tests with different densities and vertical surcharge levels, but did vary with the length of the cavity. Figure 6.13 exemplifies such differences, showing the orientation of the strain increments for two tests with different densities, surcharge stresses and cavity lengths. The lines in Figure 6.13 correspond to streamlines in the vector field, and therefore show the direction of the strain increments only, regardless of their magnitude.

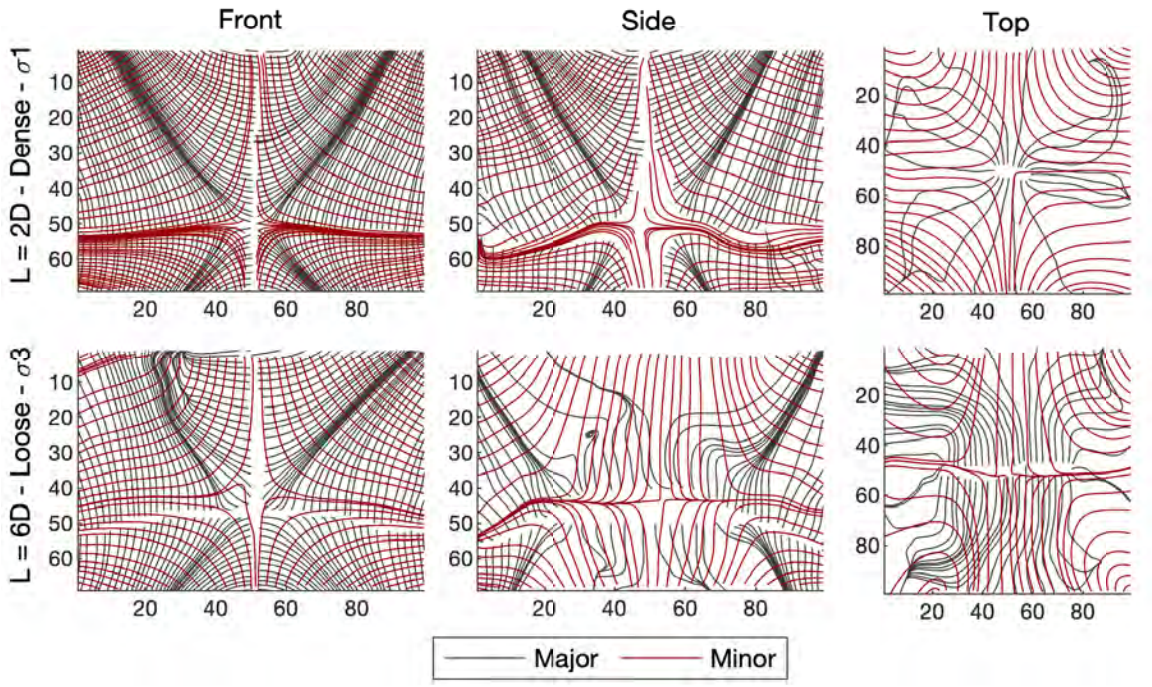

Figure 6.13: Orientation of principal strain increments. Lines correspond to streamlines following the orientation of the major and minor principal strain increments. Shown orientations correspond to the projection of the corresponding eigenvectors onto each view (front, side, top).

From the front view, we observe that the minor strain increment is horizontal on both sides of the expanding cavity, while it is vertical at the top and bottom of the cavity. Conversely, the major strain increment is parallel to the cavity at the crown, bottom and sides of the cavity. In between, the major strain increment is oriented at an angle, forming an "X" shape. It is worth noting that the shear bands are oriented in the direction of the major strain increment, while the minor strain increment is normal to it.

From the side view, we note that the orientation of the strain increments around the shorter cavity (L = 2D) is similar to that in the front view, suggesting that the failure mechanism is close to that of a spherical pressurized cavity. By contrast, the longer cavity ($L = 6D$) still exhibits the "X" shape of the major strain increment (oriented along the shear bands), but with an intermediate region in which both the major and minor strain increments are normal to the body of the cavity. From the top view, we note that the major principal strain increment is oriented normal to the body of the cavity. In the middle of the

cavity axis, the major and minor strain increments are normal to the cavity and parallel to each other, suggesting the existence of plane strain conditions on a certain region.

The stress and strain fields around the cavity are non-coaxial during early steps of plastic expansion due to the reorientation of principal stresses around the cavity [261, 262, 263]. After the soil has reached its yield point, it was shown by other authors that stress and strain become coaxial [264, 265]. Hence, we hypothesize that the orientations of strain increments shown here correspond to the orientation of the principal stresses inside the soil as well.

Previous studies on the orientation of principal stresses during arching date back to [266] who studied arching (active trapdoor) in vertical trenches or silos and proposed the orientation of the minor principal stresses forms an arch that follows a catenary shape, as described by Eq Equation 6.2. Such observations were validated by [267] who showed that the shape resembles a catenary or a circle. More recently, [268] found that the shape of the catenary arch formed above the trapdoor follows the orientation of the major principal stresses and that the peak friction angle influences the extent of that arch. Furthermore, [269] found that such arch formed along the orientations of the principal stresses develop a stress free surface that is adequately characterized by a parabola or a truncated ellipse in soils with higher friction angles.

[270] expanded the findings from [266] to the passive arching case, more relevant to our study, and found that the alignment of the principal stresses during this loading mechanism also follows the shape a catenary. Moreover, [271] studied the collapse mechanisms of shallow tunnels in cohesionless soil, and proposed that the catenary-shaped orientation of stresses appears both as downward oriented catenaries during cavity roof collapse, or as upward oriented catenaries that reach the free surface during the complete failure of the cavity, as shown in the failure mechanism exemplified in Figure 6.5 and Figure 6.13.

6.5.3 Plane Strain validation

The vast majority of analytical solutions proposed for cavity expansion problems assume plane strain conditions. We evaluate the validity of this assumption in our experiments, in which the length of the cavity was equal to two to six times the diameter. Minor and major principal strain increments are orthogonal to each other in the YZ and XY views, which is an indicator of plane strain conditions in the XZ plane (cross-section of the cavity). To quantify the deviation from plane strain conditions in the XZ plane, we define the plane strain deviation index (PSD) as follows:

$$PSD = 1 - \frac{\|U_{2D}\|}{\|U_{3D}\|} \tag{6.3}$$

where U_{2D} is the magnitude of the 2D stretch tensor ignoring the out-of-XZ-plane components, and U_{3D} is the magnitude of the 3D stretch tensor. The strain loss index L, shown as a percentage, ranges from 0% for perfect plane strain conditions, to 100% for the worst case scenario, in which all the strain occurs in the out-of-plane direction. Figure 6.14 shows the map of plane strain loss for each of the tested cavity lengths, indicating that plane strain conditions are restricted to a region around the center of the cavity, and as expected, the extent of such region increases with the total length of the cavity.

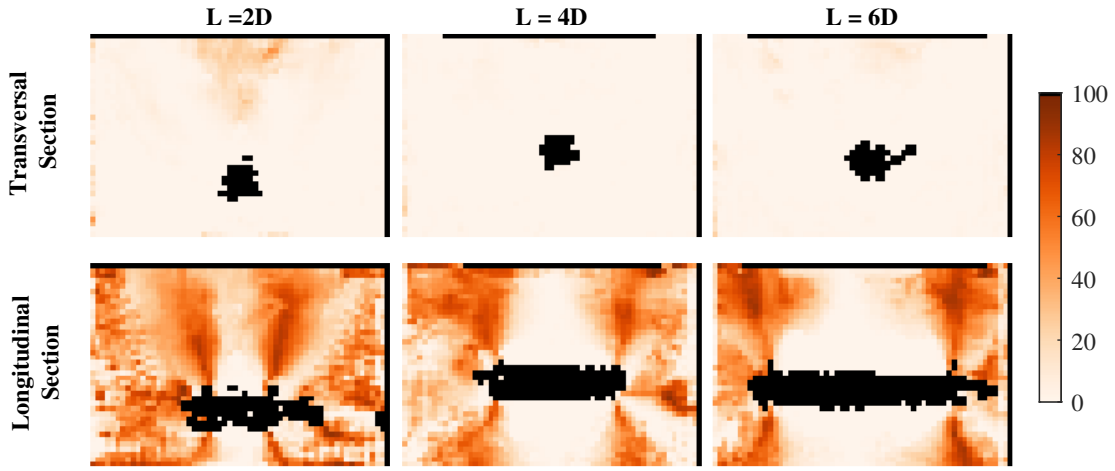

Figure 6.14: Plane strain deviation (PSD) fields. The PSD index is calculated as the proportion of strain magnitude that is lost at a given location when plane strain conditions are assumed. A value of 0 corresponds to perfect plane strain (in the XZ plane) and a maximum value of 100 corresponds to a strain field oriented completely out of the XZ plane. Shown tests correspond to dense specimens with a vertical surcharge σ_{V1}.

Furthermore, we calculated the weighted average of the PSD along the cross sections ($\overline{PSD}$) of the cavity (parallel to the XZ plane), as follows:

$$\overline{PSD} = \frac{\sum(PSD \cdot \|U_{3D}\|)}{\sum\|U_{3D}\|} \tag{6.4}$$

Figure 6.15 shows the average deviation from plane strain conditions along the axis of the cavity, for the 13 different tests, after the second inflation step. Results show that even though the deviation at the edges of the cavity is similar for all the different lengths, the region of low deviation from plane strain is larger for more slender (longer) cavities.

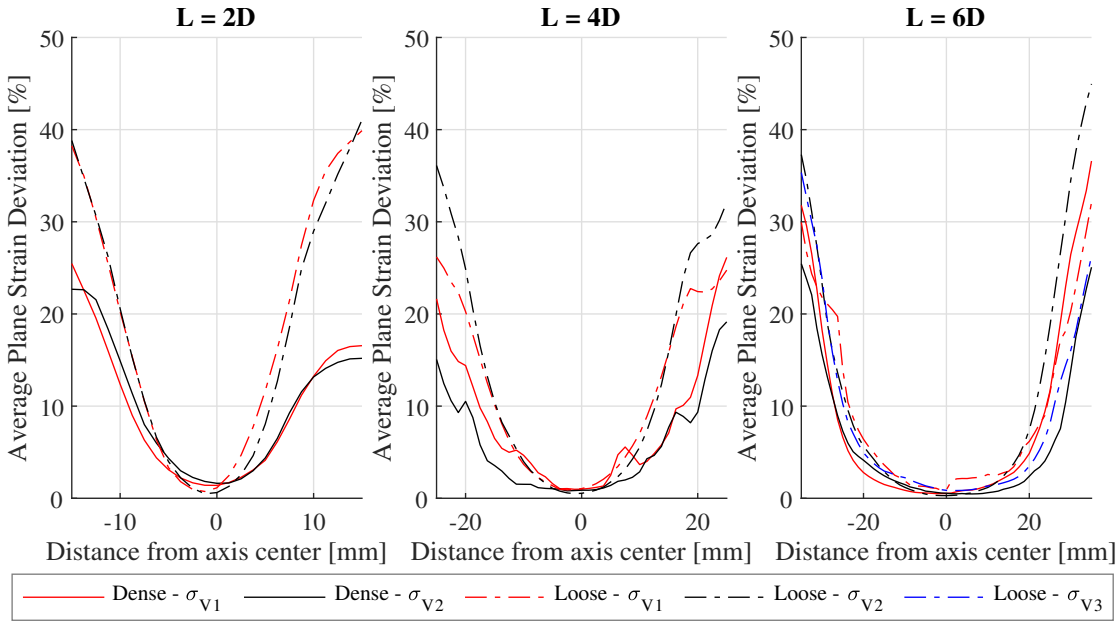

Figure 6.15: Average plane strain deviation ($\overline{PSD}$). Values are calculated as the weighted average of the PSD index along cross sections normal to the cavity axis (parallel to the XZ plane). Values are weighted by the total strain magnitude.

In order to better understand the influence of cavity length, soil density and vertical load on the extent of the region under plane strain conditions, we also calculated the percentage of the cavity length that has an average PSD below 5%. We compared that percentage for the different test variables (Figure 6.16), and we followed the evolution of that length with the inflation steps (Figure 6.17).

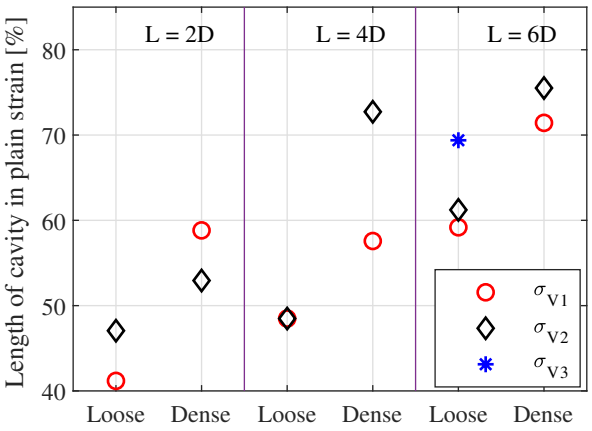

Figure 6.16: Percentage of cavity length under plane strain conditions (PS). Results from the second inflation step. The criteria adopted to set PS corresponds to the cross sections with an average plane strain deviation (PSD) under 5%.

As expected, longer cavities are closer to plane strain conditions. Yet, at least 25% of the length of the longest devices (L = 6D) did not meet plane strain conditions (Figure 6.15 and Figure 6.16). Results suggest that the higher soil density and the higher surcharge vertical stress, the more likely plane strain conditions. This observation confirms previous results on cavity deformation (Figure 6.4): a more cylindrical shape is more likely to occur at higher stress and in denser soils. Overall, pressurization promotes plane strain conditions, especially in dense soil (Figure 6.17).

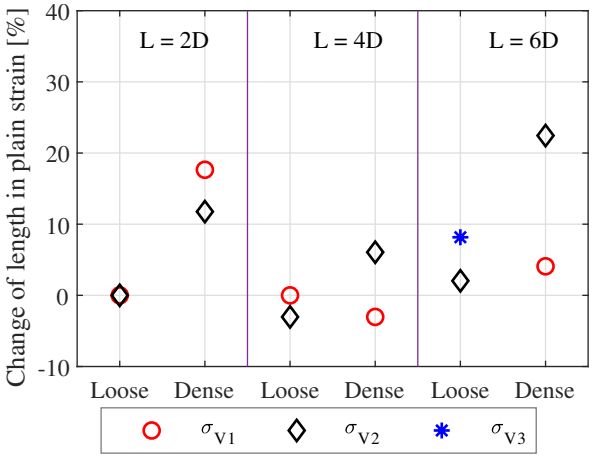

Figure 6.17: Change of cavity length under plane strain conditions (PS) between the last and first inflation steps.

6.6 Conclusions

Pressurization of a small-scale cylindrical cavity embedded in sand under geostatic stress caused a failure mechanism characterized by a concentric plastic deformation zone, typical of cavity expansion, combined with catenary-shaped shear bands typical of anchoring. The failure mechanism around shorter cavities (with length equal to twice the diameter) was closer to a paraboloid (symmetric radially) while longer cavities exhibit a plane strain region along their axis. From x-ray computed tomography images, it was noted that the soil adjacent to the cavity dilated regardless of the initial density of the soil, but inside shear bands, it either dilated in dense specimens or slightly contracted in loose specimens. The curvature of the failure planes was more concave in loose soil under low vertical stress. As density and vertical stress increased, the failure zones became closer to planes. The orientation at which the shear bands reach the surface was proportional to the dilation angle of the soil, which is a function of soil density and stress state as shown in the results.

The orientation of the principal strain increments, assumed to be coaxial with stress at large deformation stages, also aligned with a catenary shape along the longitudinal cross section of the cavity. Our study shows that the plane strain assumption currently made in state-of-the-art analytical models of cavity expansion does not hold for short cavities or tunnels at shallow depth (hence under biaxial stress) and even in the longest of the cavities tested (with a length six times larger than the diameter), plane strain conditions were only met over 75% of the cavity length. Additionally, shear strain was not negligible around the cavity, contrary to common modeling assumptions.

These results are important to the fundamental understanding of failure mechanisms at shallow depth, the prediction of failure modes and transitions of failure modes, and the design of subsurface geotechnical structures in sand, relevant to Horizontal Directional drilling (HDD), pipe uplift, micro-tunneling, and resource exploitation. Future work can extend the scope of the study to saturated or partially saturated conditions, explore the

use of well-graded soil to quantify its influence in the generations and characterization of shear bands. Additionally, changing the depth of the cavity as a testing variable could give insights in the transition phenomenon between a concentric failure zone (cavity expansion), typical of deep cavities, to the shear-band controlled mechanism found in this study.

CHAPTER 7
NUMERICAL STUDY OF SHALLOW CYLINDRICAL CAVITIES USING FEM AND MACHINE LEARNING ALGORITHMS

7.1 Aim and summary

The present chapter focuses on the application of machine learning algorithms to the problem of pressurized shallow cavities studied in chapter 6. We further analyze the problem by using finite element models that allow to automatically modify the soil parameters, and study their effect on the response of the cavity.

In the first part of this chapter, we estimate of the blowout susceptibility during drilling of horizontal cavities using machine learning algorithms, as a proof of concept. We propose the use of a predictor function, calibrated with a pool of pressure controlled numerical simulations. Blowout is predicted under given geometrical and stress conditions and based on the mechanical parameters of the soil the cavity is embedded in. We used k-means clustering (for classification) and support vector machines (SVM) to create the predictor function, and results achieved an accuracy of about 87% in predicting the blowout susceptibility. The content of this study has been published in [272] and [43].

The second part of this chapter includes the preliminary results of a study of the behavior of displacement-controlled cylindrical shallow cavities of increasing diameter. An automated routine generates models of varying depth and soil parameters and extract the distribution of stresses around the cavity wall as a function of radial displacement. We compare the obtained results to the predictions of cavity limit pressure given by empirical formulas used in the state-of-practice. In addition, we formulate a linear regression model to estimate the stress at which the soil at the cavity wall starts softening.

7.2 Introduction

The study of the response of geomaterials under complex loading conditions is an important portion of the geotechnical research and practice. Even though significant breakthroughs have been made towards the understanding and solution of many of these problems, oftentimes they become too complex to be solved analytically. Fortunately, advances in computational science and increase of computing power have made it possible to adopt numerical methods to solve these complex problems with exceptional detail and within reasonable computing time.

This same increase in computational power has generated unprecedented amounts of data, both because of the increased number of simulations and because of the level of detail achieved in these simulations [273]. The necessity to analyze and interpret all this data has deemed conventional tools as insufficient to a level such that data science has become a leading field both in scientific research and in industrial practice [274, 275]. These new (or reinvigorated) data science tools have been encased into the keyword category of machine learning, the use of which is ubiquitous to any research discipline, including geotechnical engineering and its goal of understanding the response of soils to complex loading conditions [276, 277].

In the present study, we use of machine learning, specifically supervised and unsupervised learning, as tools that can aid both scientific research, giving insights towards the understanding of controlling mechanisms; and the industry, by creating reliable and fast tools and functions that help on projects design.

One particular problem that lacks an analytical, accurate solution is that of finding the maximum internal pressure that can be applied inside the cavity before causing unconfined shear failure (blowout) of the soil surrounding the borehole during drilling. In this chapter, we focus on the problem of a pressurized cylindrical cavity embedded inside a soil subjected to anisotropic far field stresses under drained loading conditions. Even though

this problem is pertinent to several applications such as micro tunneling, in-situ testing [30, 31, 278] (pressuremeter testing under anisotropic conditions) and resource storage and withdrawal [279]; Horizontal Directional Drilling (HDD) has been the primary field of application for blowout susceptibility [37, 280, 36].

The role of the drilling fluid is crucial during every stage of the HDD process, including the boring of the pilot hole, its enlargement to the objective diameter (reaming) and the installation of the product pipe into the bored cavity (pullback) [281]. This fluid acts as a cooling agent for the drill head, suspends and transports the cuttings from the excavation front to the surface for effective drilling and it can even be responsible for rotating the drill bit [282]. Therefore, having an appropriate fluid pressure inside the borehole is paramount for HDD operation at any of its stages. However, if the minimum required pressure of the drilling fluid exceeds the maximum allowable for the soil at a given point in the borehole, mud can return inadvertently or drilling fluid can be lost into the surrounding soil, which can affect the integrity of surrounding structures (e.g. pavements, foundations), water bodies (subterranean aquifers, ponds and rivers) or cause significant surface settlements [283, 36].

The current state of the practice relies mainly on two equations to find this maximum allowable pressure value; the first is commonly known as the Delft equation [36] developed in 2001; it is the first closed form solution that was proposed to quantify the maximum allowable pressure of the drilling fluid inside the bore cavity. Nevertheless, it assumes isotropic far field stresses and cylindrical deformation of the cavity, plus it assumes the cavity is deep enough so that the cavity can be assumed to be in an infinite domain.

These assumptions are not realistic in practice, and may become non-conservative as far field stresses become anisotropic [37, 284]. The second widespread equation used to quantify the blowout pressure, known as the Queens equation [37], releases some of the assumptions made in the Delft equation, by using an elasto-plastic material and anisotropic stress conditions. Nevertheless, it is designed only for pure cohesive materials and stills

assumes that the plastic region is circular and encompasses the cavity. Several other authors have attempted to propose new, improved solutions including [285, 38, 284]. Still, the lack of consensus and limited practicality of new solutions has caused the state of the practice to still rely on the Delft and Queens equations, as can be seen on the Best practices manual from the NASTT on its latest version to this date (2017).

In this study, we create finite element (FE) models under two different loading conditions: (i) pressure controlled, where an increasing pressure is applied normal to the wall of the initially cylindrical cavity, and (ii) displacement controlled, in which the radius of the cavity is sequentially increased.

In the pressure controlled models, we fix the cavity's geometry and run the simulations using a wide range of combinations of the mechanical parameters from the soil that control its response. Then, we analyze the soil response by describing the geometry of the cavity and that of the plastic zone after the internal cavity pressurization. Geometric parameter combinations are then classified (using k-means algorithm) into 3 categories defined as low, mid and high susceptibility of blowout; finally, we make use of supervised learning (SVM) to generate a function that can predict the blowout susceptibility based on the mechanical parameters of the soil surrounding the expanding cavity. In the displacement controlled models, we vary the depth of the cavity as well as the mechanical parameters of the soil. By changing the depth of the cavity, we modify the geo-static stress state around the cavity, and with it, the distribution of stresses around the cavity as a function of the radial displacement.

7.3 Pressure controlled - blowout susceptibility

7.3.1 Finite Element Model

A finite element model is built in Abaqus to generate the source data that will later feed the classification/prediction algorithms. This model consists of a plane strain (2D) domain which approximates the cross section of the cavity since the drill length is significantly larger than the diameter of the cavity. Additionally, the existence of two planes of symmetry

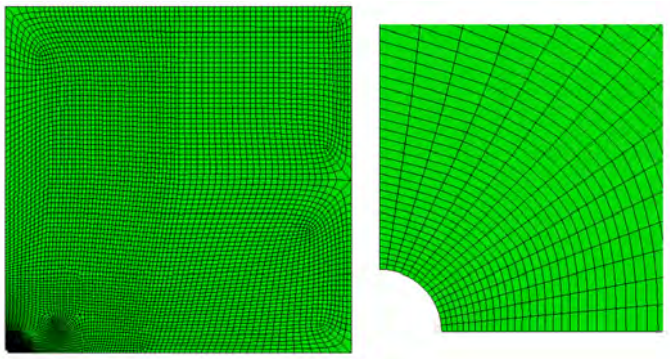

Figure 7.1: FEM domain of the simulation, the length of the edges of the square domain are 100 times larger than the cavity radius in order to avoid boundary effects. Left: full domain, Right: Close up of the cavity at the lower left corner.

allowed us to use a quarter of the domain. The square domain simulated in the following is shown in Figure Figure 7.1.

The radius of the cavity was set to $1m$ and, in order to avoid any possible boundary effects, the width (and height) of the model were set to $100m$. The corresponding symmetry boundary conditions were applied at the bottom and left edges of the domain; the right edge of the domain was fixed in order to generate a horizontal reaction to the vertical stress (imposed later). During the first step, when the far field stress is applied, the displacement of the nodes along the edge of the cavity is fixed in order to prevent unrealistic and excessive deformations. This boundary condition is later removed when the outwards radial pressure is applied.

The vertical far field stress was fixed to 200 kPa throughout the simulations and similarly, the outwards radial pressure applied at the cavity nodes was fixed to 500 kPa. The horizontal far field stress is controlled by the at-rest lateral earth pressure coefficient (K_o) which is a function of the friction angle of the material. We used the Drucker-Prager (DP) constitutive model for the elasto-plastic soil, because it ensures better convergence than the Mohr-Coulomb (MC) model.

Still, since the limits and meaning of MC resistance parameters are more broadly known than the DP ones, we chose MC parameters and used the matching plane strain DP param-

eters in the FEM model. The match between the plane strain response of the two models was obtained from Abaqus' documentation and is explained in the following equations:

$$\sin\phi = \frac{\tan\beta \cdot \sqrt{3(9-\tan^2\psi)}}{9-\tan\beta\tan\psi} \tag{7.1}$$

$$c \cdot \cos\phi = \frac{\sqrt{3(9-\tan^2\psi)}}{9-\tan\beta\tan\psi} \cdot y \tag{7.2}$$

Where ϕ, ψ and c are the Mohr-Coulomb friction angle, dilation angle and cohesion; and β and y are the DP friction angle and yield stress respectively.

7.3.2 Data generation - Pool of simulations

In order to study the blowout potential of a cavity under the loading conditions described, we run a pool of simulations with mechanical parameters varying over a range of values that are realistic for soil. Two elastic parameters are varied: the Young's modulus (E) and the Poisson's ratio (ν). The other two parameters varied thorough the tests are the Drucker-Prager friction angle (β) and dilation angles (ψ) of the material; a relatively high, constant shear yield stress (also known as Drucker-Prager cohesion) of 5 kPa is assigned to the material, since completely cohesionless materials result in early localized plastic strains around the cavity, resulting in lack of convergence and therefore incomplete data sets for the subsequent analysis. The Drucker Prager dilation angle (ψ) is set as a percentage of the Drucker Prager friction angle (β) within the range $0 < \psi < \beta$.

The present study covers a broad range of values applicable to each parameter, going from fine to coarse granular materials and from loose to dense configurations [39, 286] as shown in Table 7.1. For each one of the four mechanical parameters, five evenly spaced values within the range are set. From all the possible different parameter sets formed, a total of 625 (5^4) different combinations were simulated.

Table 7.1: Ranges of values for the soil mechanical parameters.

Parameter	Units	Lower Bound	Upper Bound
Young's Modulus (E)	kPa	10	100
Poisson Ratio (ν)	-	0.15	0.45
MC friction(ϕ)	Degrees	20	45
MC Dilation(ψ)	$\%$ of ϕ	5	95

7.3.3 Response variables - Output characterization

After the simulations finished, the elastoplastic (EP) boundary was found by interpolating the values of the plastic strain stored at each element integration points. In order to get a smooth boundary, the boundary was defined as the contour corresponding to a level curve with an interpolated plastic strain of 1e-3. The deformed cavity shape was also retrieved from the output database, following the final coordinates of the nodes along the cavity. One of the simulations corresponding to a combination of very low stiffness and resistance parameters could not converge and aborted because of excessive distortion of the nodes when the internal pressure was applied, therefore, a total of 624 scenarios were considered in our analysis.

Contrary to the common assumption that the EP boundary is ellipsoidal, EP boundary shapes varied widely, progressing towards a localized plastic region that forms a band at a given angle towards the surface. The extent and orientation of these plastic regions are reminiscent of the behavior observed, for instance, in the shear bands developed on trap door tests [287].

Once the boundaries were extracted, they were characterized by a set of shape indexes. In the case of the EP boundary, four different indexes were defined: (i) The area of the plastic region plus the cavity (calculated from the polygon that encloses the EP boundary and the center point of the cavity); (ii) The solidity of the region (a convexity metric defined as the ratio between the area of the region to the area of its convex hull – the smallest convex polygon that encloses the region), (iii) The distance between the center of the cavity to the

point along the EP boundary that is the furthest away from the cavity center; (iv) The angle from the horizontal of the line joining the cavity center to the point that is the furthest away from the cavity center.

The area of the region is a measure of the extent of the plastic zone, not taking into account its shape. The solidity of the region is a value between 0 and 1 that decreases as the shape becomes less convex, where 1 corresponds to a convex region (an ellipse section for instance). Index 3 tracks the localized development of narrow elasto-plastic bands that appear and quantifies their extension, distancing from the center of the cavity. Lastly, index 4 measures the orientation of this fingering region. If the case the EP boundary is a perfect ellipse, indexes 3 and 4 correspond to the length of the major axis and a 90° angle, respectively.

The shape of the deformed cavity was fitted to an ellipse with great accuracy in every case and its area and eccentricity were calculated to describe the cavity size and shape respectively. The eccentricity of an ellipse is defined as the ratio of the distance between the foci of the ellipse and its major axis; a value of 0 corresponds to a circle and the maximum value of 1 corresponds to the degenerate case of a line.

7.3.4 Unsupervised learning - Blowout susceptibility

From the extracted elasto-plastic boundaries, it becomes evident that the common assumption of an elliptical plastic region around the cavity is only valid when the soil is able to resist the internal pressure with little plastic deformation, nevertheless, when the soil becomes weaker (in relation to the internal pressure), the extent of the plastic region increases rapidly, deviating from an elliptical shape and causing blowout. Since the relation between the strength and deformability of the soil and the internal pressure is highly complex, and since no comprehensive analytical solution is available, we classify the elastoplastic boundaries in terms of their susceptibility to blowout.

We define three categories: low, medium and high susceptibility to blowout. Making

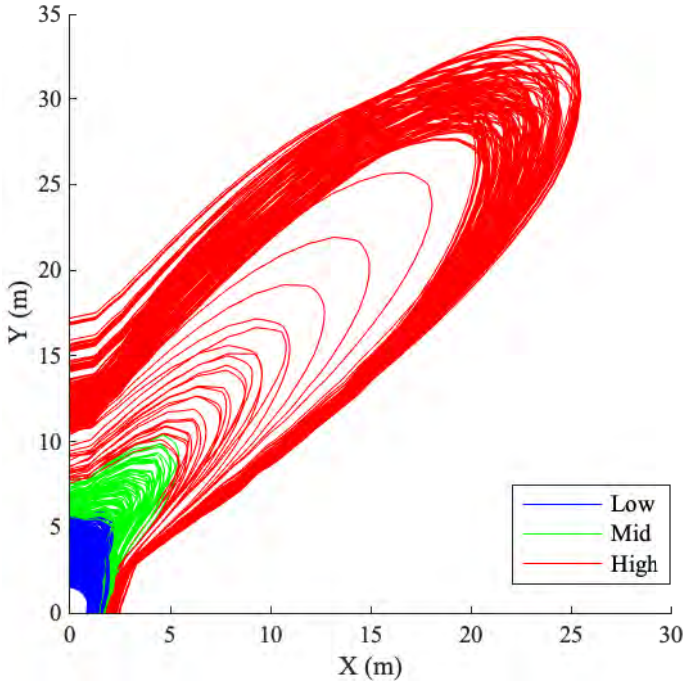

Figure 7.2: Elasto-plastic boundaries for all the numerical simulations (624) classified into susceptibility to blowout categories.

use of the geometric indexes defined for the boundaries, we use an automated classification technique, formally known as an unsupervised learning algorithm, in order to assign a category to each one of our performed simulations. We used a K-means implementation with the 4 different indexes as the features, 3 clusters (classes to classify into) and sample correlation as the similarity metric [96]. Figure 7.2 shows all the EP boundaries and the corresponding blowout susceptibility category they were classified to.

We then analyze the distribution of the shape indexes by class. A good index should be able to split clearly the different categories, meaning that it should show very little overlap between the histograms of the different classes. The frequency histograms are shown in Figure 7.3.

From these histograms, we can observe that there is a clear separation between the classes, suggesting a good choice of indexes, and confirming the expected relationship with blowout susceptibility. High susceptibility simulations show significantly higher plastic region areas (the horizontal axis in Figure 7.3 is shown in log scale) in addition to highly

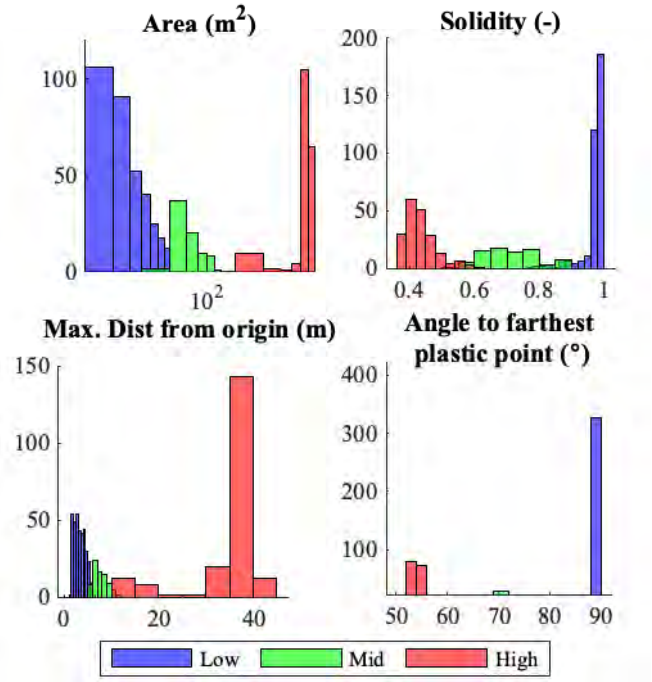

Figure 7.3: Frequency histograms by blowout susceptibility category, shown for the four EP boundary indexes: Area (m2), solidity, max distance from origin (m) and angle to farthest plastic point (degrees).

non-convex shapes, due to the generation of a preferred plastic region extending far from the cavity at a relatively constant angle from the horizontal. This "fingering" phenomenon is supported by the distance to the cavity center and the angle to the furthest point (index 4).

Conversely, low susceptibility regions are elliptical, showing the angle to the furthest point is consistently around 90 degrees (major axis), very high solidity (convex shape) and small plastic extent. The mid susceptibility region falls in between the extremes, showing the transition from an ellipse to a plastic band of preferred orientation (fingering effect).

Lastly, we use the same strategy with the deformed cavity indexes in order to test whether their distribution is directly linked to the blowout susceptibility. The results are shown in Figure 7.4.

From these results it becomes apparent that the eccentricity (sometimes referred to as the ovality) and extent of the deformed cavity are not explained by the plastic region that

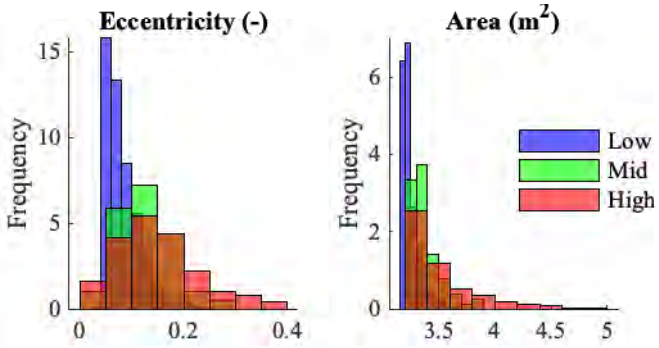

Figure 7.4: Frequency histograms by blowout susceptibility category for deformed cavity shape indexes: eccentricity and area (m^2).

surrounds it. The high overlap between categories shows that two given cavities of similar extent and shape can underlie completely different plasticity states.

7.3.5 Supervised learning - Susceptibility to soil parameters

We now make use of a supervised learning algorithm in order to generate a predictor function that returns the predicted blowout susceptibility category (classification class) based on the soil mechanical parameters (acting as predictors variables)

To this end, we tried different learning algorithms categories: classification trees, naive Bayes classifiers, support vector machines (SVM) and neural networks. After evaluating the performance of each one of the variations within these algorithms, we found out that the best performing algorithms were the neural networks and the cubic SVM, both with an accuracy between 82 and 87% for both algorithms.

Neural networks gave us a maximum accuracy of 86.7%, corresponding to a network with 10 hidden layers trained using scaled conjugate gradient backpropagation. We partitioned the data as: 70% of as the training set, and 15% for the validation and test sets each. On the other hand, the best performing SVM algorithm used a cubic boundary and 10 fold cross validation. The final accuracy of the algorithm was 82.5%.

Nevertheless, the performance of the neural network seemed to be highly variable depending on the tuning parameters and the separation of data into training and validations

Figure 7.5: Total confusion matrix for the susceptibility to blowout prediction model using a cubic SVM learning algorithm.

sets. This fact, added to the increased training time diverted us towards choosing the SVM algorithm as the most convenient, which achieved similar accuracy without the mentioned caveats. Figure 7.5 shows the confusion matrix for the predictor function gotten from the cubic SVM algorithm.

The confusion matrix shows that the class with the worst prediction accuracy is the medium susceptibility class, which is not surprising, since it acts as the transition buffer between the two extreme classes. From the misclassified cases, 2.4% correspond to low susceptibility, which overestimates the blowout potential. 4% correspond to the medium class, from which 2.9% correspond to underestimation of blowout (unsafe scenario). Finally, the high susceptibility class underestimates 6.9% of the cases, from which 5.3% correspond to samples of high blowout susceptibility classified as low.

7.3.6 Conclusions

The developed strategy showed promising results that could be implemented to assess blowout potential or any other complex phenomenon. The proposed tools build upon the best capabilities of FEM (to model complex geometries and loading scenarios), unsupervised algorithms (to find similarities and cluster large sets of data) and supervised learning algorithms (to find correlations and causality relationships that explain and predict the interactions within the data).

Nevertheless, further research is needed in order to increase the reliability of these algorithms, especially to prevent unsafe predictions that may compromise the stability of structures. It is important to mention as well that this type of tool should not be seen as a replacement to traditional, analytical methods, but rather as an extra tool that can hint and guide towards the development of new comprehensive methods.

7.4 Displacement controlled - Validation of current methods

We now study a shallow cylindrical cavity subjected to a displacement field at the cavity wall. The model is a plane strain (2D) domain that approximates the cross-section of the cavity, which is assumed to be long in relation to its diameter so that plane strain conditions hold. The symmetry of the model about the vertical axis allows us to use one half of the domain.

For each model, the geometry and properties of the material were varied randomly to generate the pool of simulations used to feed the prediction algorithms. For this reason, we used an automated script to randomly generate the set of input parameters for the simulation, build the model, run it, and then extract the output data used in the following steps.

Table 7.2: Ranges of values for the soil mechanical parameters.

Parameter	Units	Lower Bound	Upper Bound
Density (d)	kg/m^3	1600	2400
Young's Modulus (E)	MPa	30	100
Poisson Ratio (ν)	-	0.25	0.40
MC friction(ϕ)	Degrees	20	45
MC Dilation(ψ)	Degrees	12	20

7.4.1 Model input parameters

A total of 6 different input parameters are varied for each simulation, the first one is the depth of the circular cavity (H), measured from the surface to its center, while the rest are soil properties, namely: soil density (d), Young's modulus (E), Poisson's ratio (ν), Mohr-Coulomb (MC) friction angle (ϕ) and MC dilation angle (ψ). The results presented in this section are based on 500 simulations that used random combinations of such parameters.

The geometry of the half-domain is completely defined by the depth of the cavity (input parameter) and its diameter, which is fixed as $D_c = 1m$. The other dimensions of the model are determined by these dimensions: the distance from the center of the cavity to the bottom of the domain was fixed to 20 times the diameter of the cavity ($20m$), and the width of the domain was set as 1.5 times the depth of the cavity ($1.5H$).

The depth of the cavity was randomly selected as an integer value in the range $[550]$ (inclusive). Similarly, the soil properties were assigned random values, independent from each other, between reasonable parameters for each one of them, as shown in Table 7.2.

A value of MC cohesion (c) corresponding to $5kPa$ was added to every simulation for stability of the model. We used the Drucker-Prager (DP) constitutive model for the elasto-plastic soil, because it ensures better convergence than the more popular Mohr-Coulomb (MC) model. Still, since the limits and meaning of MC resistance parameters are more broadly known than the DP ones, we chose MC parameters and used the matching plane strain DP parameters in the FEM model. The match between the plane strain response of

the two models was obtained from Abaqus' documentation and is explained in Equation 7.1 and Equation 7.2.

7.4.2 FEM model and output extraction

Once the geometry and soil properties are defined by the input parameters, the built-in Abaqus scripting capabilities are used to build the model, including the loading steps, boundary conditions and mesh. The initial geo-static stress conditions are applied in the first step of the model, with an implicit model (Abaqus Standard), while the expansion step was modelled using a quasi-static dynamic step.

First, the geometry of the problem is set, soil properties are assigned, as well as the initial boundary conditions and the geo-static stress state. The second step releases the boundary conditions around the cavity and free surface at the top of the model and imposes a displacement-controlled expansion of the cavity using a smooth amplitude ('s' shape), up to a maximum radial displacement equal to three times the initial diameter.

The objective of the modeling effort is to simulate the expansion of the cavity as a response to a mechanically induced radial displacement, therefore, only the reaction forces at the cavity boundary are extracted from the model output. We monitor the reactions at the 37 equally spaced nodes around the cavity wall at every expansion increment (neighboring nodes are separated by an angle of 5 degrees).

7.4.3 Stress distribution around the cavity

From the original output of nodal forces around the cavity, we find the equivalent radial stress by dividing the resulting force by its afferent area. Then, we have a cavity response at each radial orientation, starting from the horizontal (defined as the angle $\alpha = 0$) to $\alpha = 90$ at the crown of the cavity (top), and to $\alpha = -90$ at the invert (bottom), at 5 degrees intervals. The radial displacement at each increment is normalized by the initial radius of the cavity ($R_c = 0.5m$). Figure 7.6 shows an example of the obtained stress distributions

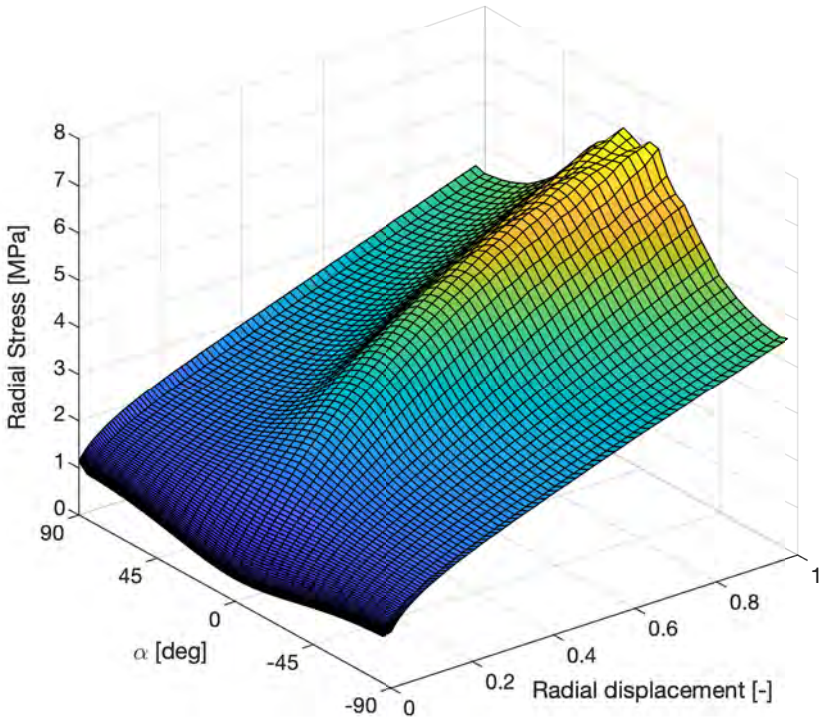

Figure 7.6: Radial stress distribution around the cavity wall. The example presented corresponds to a cavity at a depth $H = 42m$, embedded inside a soil with the properties: $\phi = 31°$, $\psi = 12°$, $E = 30.5MPa$, $\nu = 0.38$, $d = 2380kg/m^3$.

around the cavity

The distribution obtained from relatively deep cavities, such as the example shown in Figure 7.6, exhibits a symmetry about the horizontal, e.g. the radial stress at an angle β is similar to that found at an angle $-\beta$. This fact supports the common assumption of biaxial stress as a simplification of the actual geo-static stress field, which causes stress variability along the vertical axis of the cavity [243]. In the Delft equation [36], the stress field is assumed to be isotropic ($K_o = 1$). Even though it is well known that this assumption is unrealistic at any stress level [288], this equation remains the accepted method to calculate the limiting pressure of the cavities in horizontal directional drills (HDD), due to the lack of practical predictive methods for frictional materials.

Then, we compare the limit pressure from the Delft equation (P_{lim}) to the stress that causes a region around the cavity to start softening. Such softening stress is found from

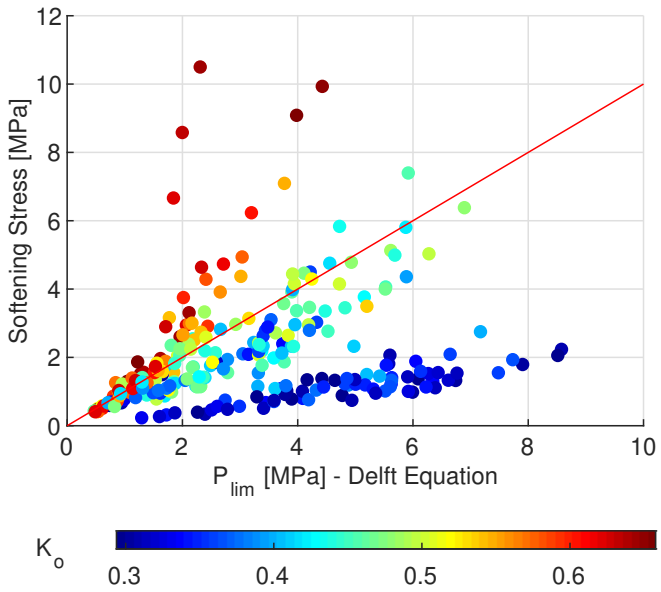

Figure 7.7: Comparison between the limit pressure (P_{lim}) from the Delft equation and the obtained softening stress from the numerical models. The color code indicates the value of the earth pressure coefficient at rest K_o for each data point.

the stress distribution around the cavity (see Figure 7.6). For each stress orientation, we found the point at which the material softens, e.g. the point at which an increase of radial increment causes a decrease in the reaction force/stress for the first time. The smallest stress at which softening occurs is defined as the softening stress, and its value compared against the prediction of the Delft equation. Figure 7.7 shows the comparison between the values for the obtained simulations, in which the color of the data points maps to the coefficient of earth pressure at rest K_o, which is a function of the friction angle of the material.

Results from the Figure 7.7 show that the Delft equation fails to predict the limit pressure values, yielding increasingly unsafe predictions as the far field stress conditions diverge from an isotropic state, e.g. as K_o goes farther away from 1.

7.4.4 Prediction of softening stress from input parameters

The current section is part of the ongoing and future work of this thesis and only intends to assess the feasibility of using machine learning models to train a model that can predict the response of the cavity. From the set of input parameters of the simulations (predictors) and

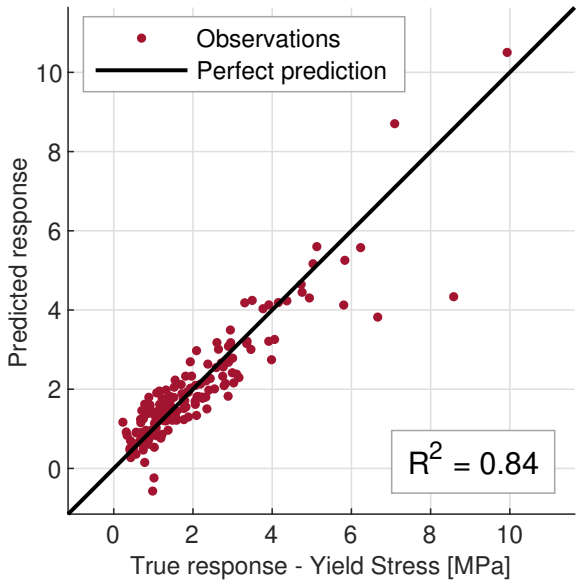

Figure 7.8: True vs. Predicted response. Linear model that fits the softening stress of the cavity as a linear combination of the set of predictors, e.g. the mechanical parameters of the soil and the geometry (depth) of the cavity.

the calculated softening stress (response variable), we fit a linear model and check if that model can predict the response of the cavity.

The angle parameters (ψ and ϕ) are transformed into their sinus value because they usually appear in that form in analytical expressions, e.g. in the expression of K_o. In the same way, the vertical stress at the center of the cavity $\sigma_v = g \cdot Hd$ is added as a predictor. Given that the cohesion of the material is the same for all the simulations, it has no predictive power and is removed from the set of predictors.

A robust linear model is fitted to the data, in order to find a first correlation between the properties of the soil and the response of the cavity. The fitted model has a coefficient of determination $R^2 = 0.84$, meaning that 84% of the variability of the softening stress is explained by a linear combination of the used predictors. Figure 7.8 shows the correlation between the true response e.g. the softening stress of the cavity, and its predicted value from the linear model for the tested observations (simulations).

Results suggest that even a relatively simple linear model can achieve a good prediction of the softening stress of these cavities. Ongoing work on this topic focuses on developing

more accurate models, understanding the correlations between variables of the model, and modeling soil response from a physically-based point of view, in order to better inform the development of new, more accurate semi-analytical solutions.

7.4.5 Conclusions

The results obtained from this section show that the current methods to estimate the response of shallow cavities are inaccurate as a result of the assumptions and simplifications needed to achieve practical closed-form solutions. In the case of HDD and the Delft equation, we showed that the inaccuracy of the method is related to the absence of isotropic stress conditions around the cavity. The preliminary work shown in this section suggests that it is feasible and effective to use machine learning models to predict the response of cavities accurately, and that it is possible to use machine learning as a readily available verification tool during for project design. More importantly, these models can inform the development of more accurate analytical and semi-analytical solutions by shedding light on the influence of the different soil parameters.

Future work will further explore the effect of cavity depth, since results from chapter 6 suggest that there is a transition from the symmetric failure mechanism seen around deep cavities, to a failure mechanism that is influenced by the vicinity of the free surface. In addition, future work on the topic will include the use of a constitutive model that better captures the behavior of the soil, e.g. plastic hardening/softening. In addition, conventional FEM is limited to relatively small deformations in the domain, therefore, next steps should consider the use of large deformation methods such as MPM (material point method) or at least, the use of re-meshing during the simulations.

CHAPTER 8
HORIZONTAL DIRECTIONAL DRILLING (HDD) ALIGNMENT OPTIMIZATION USING ANT COLONY OPTIMIZATION

8.1 Aim and summary

The present chapter focus on the development of an optimization algorithm for Horizontal Directional Drilling alignments. Horizontal Directional Drilling (HDD) is a trenchless method that consists in drilling an inclined and curved bore from an entry point to an exit point. In practice, HDD is designed iteratively by trial and error, to minimize the cost under geometric and mechanical constraints. We optimize the drill path with continuous implementations of an Ant Colony Optimization (ACO) algorithm that sets the depth of the alignment and its entry and exit angles as the design parameters to optimize, to ensure minimal drill path length (cost), avoid collapse or instability (mechanical constraints) and remain in the construction domain (geometric constraint).

Then, we compare the ACO results to the drill paths designed in practice in two different scenarios: one in which the entry and exit points are fixed, and one in which the geometry of the central segment is constrained. Results show that ACO can be used to automate the otherwise time-consuming design process while minimizing the drill path length and the costs associated to it. The content of this chapter was published in [44].

8.2 Introduction

Horizontal directional drilling (HDD) is used to drill cavities through geomaterials at relatively shallow depths. The basic sequence of steps of an HDD installation consists in a preliminary design, followed by the drilling of a pilot bore from the drilling rig (entry point) to the exit point following the geometry of the designed path. Then, the pilot hole is

reamed, i.e., it is sequentially enlarged by changing the drilling head used for the pilot bore with a reamer, which concentrically increases the cavity along the pilot alignment. Finally, the pullback step consists in the installation of the product or casing pipe, which is pulled through the reamed borehole, usually from the exit point towards the entry. Compared to conventional cut and cover HDD presents advantages in cost, environmental impact, land use and project timeline [289, 290]. HDD is also advantageous compared to other trenchless methods. For instance, HDD is usually less expensive than micro-tunneling (MT), despite the fact that HDD alignments are typically longer than MT alignments. This is because HDD is done with less specialized equipment and does not require entry/exit shafts to reach the drill depth. While auger drills are easily infiltrated by pore water under high groundwater tables, HDD performs well under high porewater pressure. Additionally, the jacking loads and torsional stresses on the auger flight usually limit the maximum drill length of auger borings [289, 291, 290]. HDD is exempt of that limitation.

Since its first implementation in 1971 in a natural gas installation crossing the Pajaro river in Watsonville California [292], HDD has been successfully adapted to complex geological conditions and geometrical constraints. To date, the longest installation with a single drill rig crossed the Qin river in Jiaozuo, Henan province in China with a total drill path length of 1.75km. Installations with two drill rigs achieved a maximum drill length of 3.3km – for example, the gas pipeline installation across the Yangtze river in China [292].

Despite its increasing popularity in the engineering practice [289], many fundamental mechanisms pertaining to the design of HDD remain unknown. Current research aims to increase the alignment and maximize the cost efficiency of HDD [293, 294, 37]. Mechanical evaluations of HDD focus either on borehole instability or on pipe integrity [289]. By contrast, no method exists to optimize the drill path, which is still determined iteratively to this date: an initial estimation of the alignment is tested and gradually adjusted until a path that complies with the design criteria has been found [291]. To overcome this limitation, we explain and test a numerical method that sequentially calculates the alignment of the drill

path to satisfy mechanical design constraints (here, borehole stability and pipe integrity) while minimizing the installation cost (here, the total drill length). We first present a design method based on geometrical constraints (e.g., constructability angles, segment lengths and depth of cover from the ground surface and obstacles inside the soil) in section 8.3. Then we explain a design method based on mechanical considerations (i.e., borehole stability and pipe integrity) in section 8.4. section 8.5 summarizes the general framework of Ant Colony Optimization (ACO) and explains the current implementation of ACO applied to the specific problem of HDD alignment design (ACO-HDD). The ACO-HDD method is tested for two design scenarios in section 8.6. section 8.7 presents the conclusions of this chapter.

8.3 HDD alignment design

Approximate locations for the entry and exit points of the drilling path are decided after completing a preliminary study that also assesses the relevance of HDD for the project at stake. Then the geometry of the HDD is designed to fit the drill path within the assigned domain. The mechanical design is the last step, to test borehole stability and pipe integrity during the whole construction. If the initial design does not comply with the assigned constraints, it is iteratively adjusted.

8.3.1 Geometric design

The geometric alignment considered in here consists of at least 5 segments, starting with an entry tangent from the rig side, followed by a curved segment that reaches the central portion of the alignment. This central portion consists of at least one straight segment, but more complex alignments may include extra straight and curved portions (with vertical, horizontal or compound curvature). The central portion is typically designed first, followed by the adjacent segment and the exit tangent. The radius of curvature of the curved segments is chosen so as to prevent excessive bending stresses as the pipe is pulled back

from the surface through the curved segments, and thus depends on the material properties and diameter of the pipe. In this study, the radius of curvature is fixed. Therefore, once the entry/exit angles of the vertical projection of the curve are set (from the end points of the adjacent segments), the vertical projection of the curve (δy in Figure 8.1) is fully determined. In the following, we use the terms "vertical curve" and "vertical projection of the curve" interchangeably.

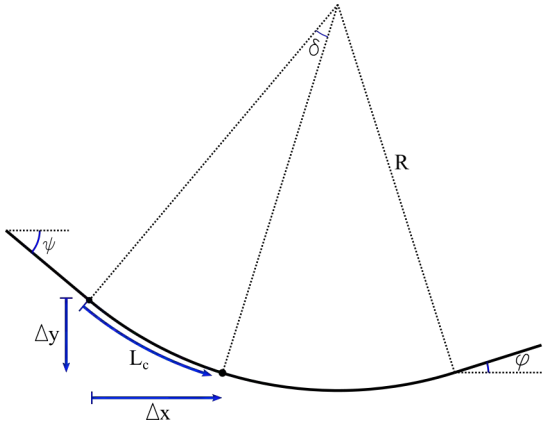

Figure 8.1: Vertical projection of the HDD curve diagram, with parametrization along the x direction.

The parametrization of the horizontal projection of the curve (Δx in Figure 8.1) as a function of the stationing of the project is found according to the following equation (Equation 8.1):

$$\begin{aligned} S &= sign(\phi - \psi) \\ \delta &= sin^{-1}(\psi + S(\frac{\Delta x}{R}))) \\ \Delta y &= S \cdot R(cos(\psi) - cos(\delta)) \\ L_c &= R \cdot \delta \end{aligned} \tag{8.1}$$

Good practices recommend the inclusion of straight segments of a certain minimum length preceding and following curves for constructability. Figure 8.2 shows an example

of an HDD geometric alignment in plan view (top view, plane XY) and in profile view, i.e. following the direction of the stationing of the alignment in the horizontal axis and the elevation in the vertical axis.

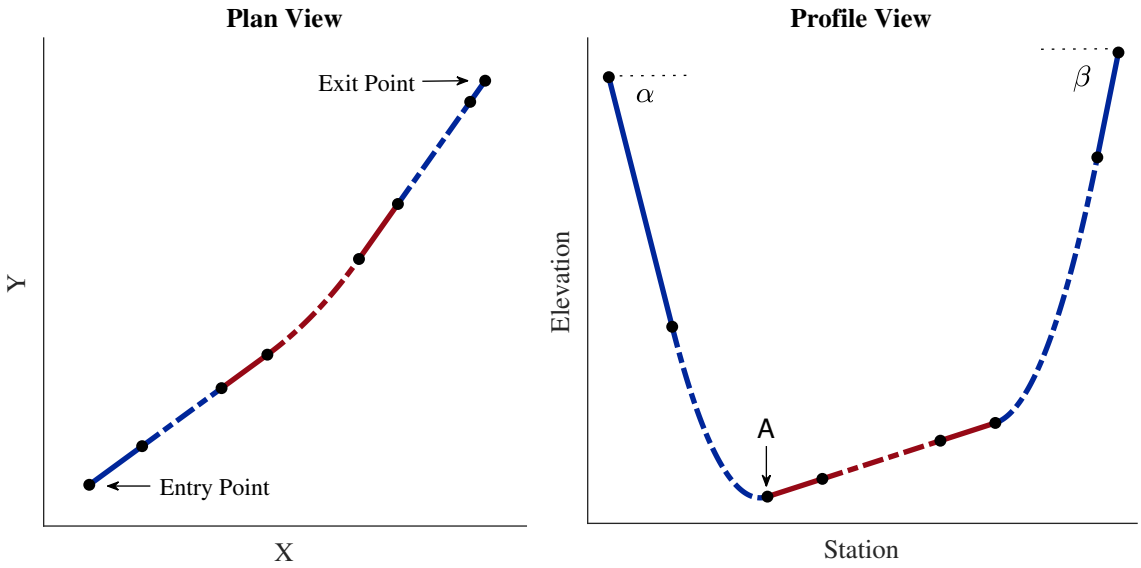

Figure 8.2: Example of a typical HDD alignment. Left: Plan (top, XY plane) view Right: Profile (Stationing vs Elevation) view. Solid lines correspond to straight segments while dashed lines correspond to curves. Central portion of the alignment highlighted in red.

The alignment design is optimized by adjusting three parameters: the elevation H of point A in the central portion of the alignment, the entry angle (α) and the exit angle (β) which control the entry and exit tangents and their neighboring vertical curves. The range of variation of α and β is constrained by constructability limitations and is usually between of 5 and 18 degrees. We adopt the convention that the drill progresses from the rig side, located on the left side of the diagrams, to the exit side located on the right; angles are measured in reference to the horizontal line and are counted positive when going counter-clockwise – therefore, α and β are positive and negative respectively.

A geometrically viable alignment is such that: (i) It can still fit within the fixed bounds of the HDD drill. (ii) It yields a minimum depth of cover along the alignment that is at least equal to the minimum allowable, in order to prevent surface disturbance and damage to the product pipe. Usually the minimum depth is between 5 and 15ft, depending on the

application, soil/rock conditions, surrounding foundations and existing utility networks. (iii) The length of every straight segment is at least equal to the minimum defined for constructability. This minimum usually corresponds to one to three times the length of the drilling rods used for drilling; in some cases, this minimum can be set to zero, creating alignments with no transition between curves.

Lastly, the parameter H varies in a range that depends on the other geometric parameters. The minimum value of H, i.e. the lowest possible elevation of point A, and therefore the deepest drill path, is obtained by setting the entry and exit angles to their maximum magnitude and setting the entry and exit points as far away from each other as possible. Conversely, the maximum value of H is the shallowest path that still complies with the depth of cover constraints. Once the geometrical constraints have been cleared, the mechanical viability of the given alignment is tested.

8.4 Mechanical Design

8.4.1 Borehole Stability

The drilling fluid plays a fundamental role in every step of the HDD process: pilot drilling, reaming and pullback. The continuous flow of drilling mud cools down the drill head and the reamers, transports the cuttings to the exit pit for effective drilling and reaming, provides lubrication for the tooling and product pipe inside the borehole, seals fractures and high permeability paths that can lead to inadvertent returns to surrounding formations, improves the stability of the cavity and thus prevents collapse [281, 295, 296, 288].

The drilling fluid is a non-Newtonian fluid, usually made from a mix of water, bentonite and extra additives. An external pump circulates the fluid from outside the borehole through the drilling rods and flows back to the free surface through the annular space between the drill rods and the borehole or between the product pipe and the borehole (during pullback) following the path of least resistance. The minimum drilling pressure (MDP) necessary to induce re-circulation of the drilling mud is equal to the sum of the difference of pressure

head between the pumping station and the drill path and of the fluidic drag developed in the annular space (controlled by the suspended cuttings and by the viscosity and yield point of the fluid).

The injection of the pressurized drilling fluid induces a deformation of the borehole. This phenomenon was modeled using the theory of cylindrical cavity expansion under different scenarios [296, 36, 37]. The so-called Delft equation [36] assumes that the far field stress around the cavity is isotropic, and that the cavity wall and the elasto-plastic boundary are circular. It is assumed that cavity shear failure, known as blowout, occurs as the internal pressure reaches the limit pressure of the cavity, causing large radial displacements and plastic expansion. The Queen's equation [37] releases the assumption of isotropic far field stresses, considering biaxial stresses, but still assumes a circular elasto-plastic boundary around the region. The formulation assumes that shear failure develops either at the top (crown) or at the side of the cavity, simplifying the analysis. The development of more accurate analytical solutions to this problem remains an active field of research to this day. Despite its simplifying assumptions, the Delft equation remains the most commonly accepted formulation and will therefore be adopted in the present study to find the maximum allowable pressure (MAP) of the drilling fluid before blowout is triggered.

8.4.2 Product pipe Integrity

The second step of the mechanical analysis aims to calculate the stresses (and subsequent strains) in the pipe as it is pulled back from the exit to the entry point of the drill alignment. External loads on the product pipe are:

1. The tensile force due to the pull operation from the rig. The required force to pull the pipe is a function of: the weight component of the pipe in the pull direction, the friction from the contact between the pipe and the borehole and/or the ground surface, the fluidic drag arising from the shear resistance of the viscous drilling fluid around the pipe and the extra frictional stresses due to increased contact forces on

the pipe as the pipe is bent around curved segments.

2. The bending moment applied as the pipe is pulled through curves. Bending induces extra tensile stress in the outermost fibers of the pipe, therefore the critical stress condition must consider the combined tensile stress from tension and bending of the pipe. Depending on the material that makes the pipe, the way to calculate stresses due to bending varies; ductile iron pipes usually include flexible joints which dissipate the bending stress around curves, while steel pipes behave as rigid beams when bent around curves; the resulting increase in friction from the contact is calculated as described in [297, 37]. Lastly, plastic pipes (HDPE or FPVC) are usually considered flexible and thus the increase in tensile forces is modeled using the capstan effect [298].

In addition to testing the pipe for tensile failure, it is necessary to check that the pipe remains stable, i.e., that it does not buckle. In order to avoid excessive friction between the pipe and the borehole and thus decrease the MDP, the borehole is usually considerably larger than the product pipe. Therefore, the pipe rests inside the borehole under unconstrained conditions sustaining the pressure head from the column of drilling fluid and possibly some earth loading. Buckling calculations are thus intended to prevent possible collapse or excessive oval deformation, which may alter the installation of conduits inside the product pipe [291].

Once the individual mechanisms affecting the integrity of the pipe are quantified, the deformation of the product pipe under the combination of loads is compared to the maximum recommended deflection and the pipe stress is compared to the pipe material strength parameters. Factors of safety depend on the pipe material and manufacture [291].

8.5 Ant Colony Optimization (ACO)

8.5.1 Background and literature review

Physical studies have shown the outstanding capabilities of ants to forage complex domains without any means of direct communication between them [299]. Ant colonies display a very effective swarm behavior based on the deposition of a chemical substance, which allows the whole colony to sequentially find the most efficient way to exploit the resources of a domain. Inspired by this behavior, several authors created ant-inspired computational optimization algorithms, which were then grouped into a general framework known as Ant Colony Optimization (ACO). ACO is a family of meta-heuristic algorithms that solve complex combinatorial problems [300] by using a computational swarm intelligence, where individual entities contribute to the global, collective exploration knowledge of the domain, sequentially improving the solutions to the tested problem, to converge to an optimal.

The common framework for ant inspired optimization was initially proposed for solving discrete combinatorial optimization problems and was applied to the common benchmark traveling salesman problem [300]. Other early applications of ACO include routing in communication networks, quadratic assignment problems and job-scheduling [301]. Since the initial algorithms proposed in the early 90's, the capabilities and applications of ACO have advanced significantly, being now recognized as an effective and flexible technique with a vast number of different implementations suitable for virtually any optimization conditions. Comprehensive reviews were presented in [302, 300].

Applications specific to civil engineering have focused mainly on transportation, for traffic routing more so than infrastructure design, for example for train traffic and bus network design [303, 211]. In this study we make use of ACO to find the optimal combination of geometric parameters that minimizes the total length of the HDD drilling path. We use the ACO implementation described in [304], which can be used for mixed variable problems, including discrete (ordinal and categorical) and continuous variables. In the case

studies presented below, we only use continuous variables. The implementation can be improved to include discrete variables, using the same base algorithm.

8.5.2 Algorithm principle

The basic principle of the algorithm is inspired by the foraging strategies employed by an ant colony that is inside a complex maze with numerous multiply-connected paths. Initially, the colony has no prior knowledge of the domain. Therefore, at first, individual ants explore the maze randomly. Each exploration yields a certain path. At the end of a given path, the ant may find a food source. A score is assigned to each path, in proportion to the quality and amount of the food source associated with the path. This score can be understood as a concentration of pheromone. This pheromone is deposited along the path in a process known as stigmergy, which is an indirect form of communication between ants based on a modification of the environment, which becomes the collective knowledge and memory of the colony [197].

The scores assigned to paths explored bias subsequent groups of ants exploring the same maze, which are then attracted to segments of increased pheromone concentration. Eventually, a set of optimal paths yield the best routes of domain exploitation for the colony. The intrinsic randomness of the process, in addition to the dissipation of pheromone concentration over time, ensure an appropriate exploration trade-off, where exploration is always encouraged but convergence is eventually achieved.

8.5.3 Algorithm implementation

Here, an ant corresponds to a combination of parameter values that results in a path, i.e., an alignment to be tested. We implemented an HDD ACO algorithm and tested it for two different design scenarios. The first one considers an alignment with fixed entry and exit points and a single straight segment in the central portion, with no horizontal curvature. The plan view of such alignment is a straight line connecting the entry and exit points and

remains unchanged by the algorithm. The profile view changes as a function of the input parameters. Figure 8.3, Case 1 shows two possible solutions generated by the algorithm in this scenario; the entry and exit points remain fixed for both solutions while the entry/exit angles and H vary, changing the length of the segments of the alignment with them.

The second scenario considers that the plan view of the central portion of the alignment is fixed and that its profile view can only change in elevation, while the location of the entry and exit points is flexible, as shown in Figure 8.3, Case 2. This scenario allows modeling alignments where the azimuth of the entry and exit tangents is different and/or the central portion must go around obstacles. The central portion can then include an arbitrary number of segments that can be either straight or curved. The elevation of the central segments is controlled by the variable H, used as the reference for the whole central portion. The entry and exit points of the alignment vary within acceptable ranges defined by the user.

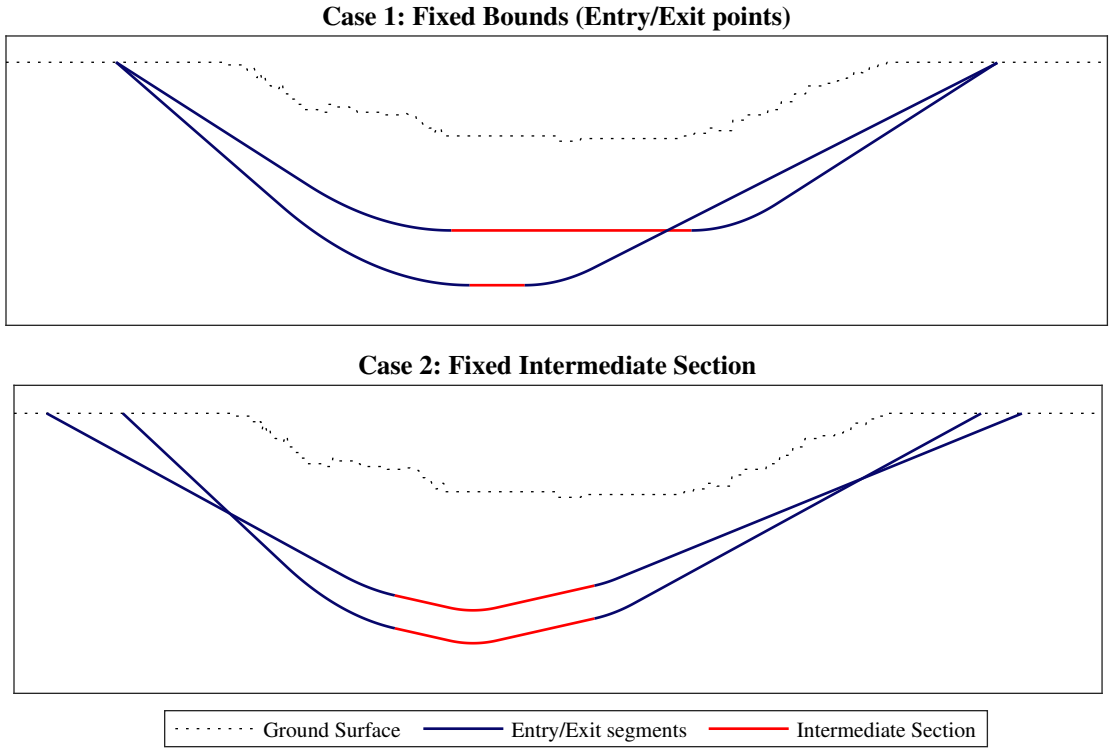

Figure 8.3: Implemented geometric scenarios. Each case shows an example of two different possible ACO solutions. Case 1 has a fixed location of entry/exit points with a variable intermediate straight segment. Case 2 shows an intermediate section configuration with a fixed geometry (curvature) but variable elevation and variable entry/exit points location.

The ACO algorithm for HDD design is implemented in two main steps. First, in the initialization of the algorithm (see section subsection 8.5.4), the domain of exploration is defined. In our case, the intervals in which the design parameters range are set. Secondly, in the optimization step (see section subsection 8.5.5), the knowledge of the domain is improved iteratively, therefore producing better solutions.

Before the initialization of the algorithm, the ranges of variation of the optimization variables are set and the size (n) of the solution ledger is fixed. Then, during the initialization step, n initial solutions are found from simple random combinations of the variables. Then the optimization algorithm runs a number (nG) of simulation batches, named generations. An ant is defined as a stochastic combination of the optimization variables (α, β and H); after an ant is constructed by the algorithm, it is evaluated. If the ant satisfies the geometric and mechanical constraints described in subsection 8.3.1 and section 8.4, and if it yields a valid design, then the ant becomes a solution, and it is stored in the ledger by the algorithm. The process is stopped when nG generations have been simulated or when a convergence flag has been triggered. The convergence flag is usually set to stop when the difference between the output of two consecutive generations is below a set threshold.

Algorithm 1 shows the outline of the implementation.

Algorithm 1: ACO-HDD Implementation

Initialization (subsection 8.5.4)

1 - Set variable ranges;

2- Find initial n viable ants (randomly);

Optimization (subsection 8.5.5)

while *Termination criterion is not satisfied* **do**

    1 - Set variable ranges;

    2- Find initial n viable ants (randomly);

    while *less than k new solutions have been found* **do**

        Construct Ant;

        Evaluate Geometry and Mechanical viability;

        If Ant is successful, then: store as solution;

    end

end

8.5.4 Initialization

Step 1 – Variables range: The first stage of the initialization consists in finding valid ranges of variation of the optimization variables. To do so, a set of project specific parameters are needed: the profile of the ground surface elevation along the HDD path, the location and geometry of any obstacles that must be avoided by the algorithm, the characteristics of the drill equipment, which partially define the range of the entry and exit angles and the minimum segment length, and other project specific parameters including the radius of curvature, usually a function of the product pipe to be installed, and the minimum depth of cover. The elevation profile and the location of obstacles along the drill path control the maximum elevation of the alignment. Once this maximum H is found, the corresponding α_{max} and β_{min} angles necessary to reach that value of H are found geometrically and ad-

justed as necessary to fit in the initial user-defined ranges. Conversely, the steepest angles α and β yield the lowest possible elevation of the alignment, therefore setting the minimum value of H.

Step 2 – Initial ants: After finding the valid ranges of variation for each one of the optimization variables, we find the initial n solutions (viable ants) that will start the algorithm. To do so, random combinations of parameters (ants) are tested, assuming that all the parameters follow a uniform probability distribution within their respective ranges. The process stops when n successful ants have been found, and their information, e.g. their values of H, α and β and the resulting total length of the alignment stored on a ledger. This solution ledger is an archive that stores the colony's knowledge about the domain and that is updated during the optimization stage.

8.5.5 Optimization

The optimization algorithm repeats a subroutine over a loop until a convergence flag is reached or a fixed number of generations has been computed. Such subroutine is detailed below.

Step 1 – Update/Sort solution ledger: At every generation, the algorithm ranks the solutions in the ledger according to their quality or score. In our implementation, we define the score of a solution as the inverse of the total alignment length, therefore the best solution is the one that yields the shortest length while complying with all the geometric and mechanical constraints. Once the solutions are ranked from 1 to n, a weight (wj) is associated to each solution j using a Gaussian function, as follows (Eq Equation 8.2).

$$w_j = \frac{1}{2qk\sqrt{2\pi}} \cdot exp(\frac{-(rank(j)-1)^2}{2q^2k^2}) \tag{8.2}$$

Where rank(j) is the rank of the solution j, and q is a user defined parameter of the algorithm. The probability of choosing a given solution (j) to create a new ant (p_j) is

calculated using the following equation (Eq Equation 8.3):

$$p_j = \frac{w_j}{\sum_i^k w_i} \tag{8.3}$$

Values of q close to zero result in a highly skewed probability towards the higher-ranked solutions against lower-ranked solutions; conversely, very large values of q result in a uniform probability to choose any of the solutions in the ledger. The probability distribution of the solutions is also a function of n. Once the solutions ledger has been updated for the given generation, new ants are created and tested until a fixed number of (k) new solutions have been found.

Step 2 – Generation ants: Once the ledger has been updated, and with it the probability distribution of the current solutions in the ledger, we start a secondary loop which runs until we have found k new viable ants. The process of constructing a new ant (i.e., a combination of variable values) starts with the selection of a solution s from the ledger based on the discrete probability distribution described above. Then, a new value is independently assigned to each variable. A new value of variable (v) from solution (s) is sampled from a normal probability distribution with mean (μ and standard deviation (σ as shown in the following probability density distribution (Eq Equation 8.4).

$$pdf(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot exp(\frac{-(x-\mu)^2}{2\mu^2}) \tag{8.4}$$

Where (μ is taken as the value of variable v in the chosen solution s, and the standard deviation (σ the standard deviation is a measure of the dispersion of the variable across the current solutions in the ledger; It is calculated as follows (Eq Equation 8.5).

$$\sigma = \frac{\varepsilon}{n-1} \cdot \sum_{i=1}^{n} |v_s - v_i| \tag{8.5}$$

The parameter epsilon (ϵ) is a user-defined parameter that affects the convergence speed of the algorithm; high values of epsilon result in high standard deviations, which slow down

convergence and vice-versa. Similar to the parameter q, the choice of epsilon depends on the application and the desired runtime of the algorithm; fast convergence may result in unsatisfactory exploration of the domain.

Once an ant has been constructed, its viability is tested in reference to the geometric and mechanical constraints of the problem. In the first case, with fixed entry/exit points, the horizontal length of the entry and exit tangents and vertical curves is completely defined by the design parameters (α, β and H), and subsequently, the length of the central segment is fully determined. The alignment is geometrically acceptable if the length of each of the straight segments exceeds the minimum segment length and the minimum depth of cover is maintained along the alignment.

In the second case, the curvature and x and y coordinates of the intermediate segment is fixed, but the entry and exit points locations change as a function of the optimization parameters (α, β and H). Therefore, in addition to checking that the length of the segments is above the minimum, it is necessary to check that the entry and exit points fall within their allowable ranges (defined by the user). Figure 8.3 shows examples of drill paths for both cases.

Once it was verified that the alignment satisfies the geometric constraints of the problem, the mechanical stability of the borehole is checked for pilot drilling and pullback; the MAP (maximum allowable pressure before blowout is triggered) and MDP (minimum drilling pressure to induce fluid re-circulation) are compared every 0.3m ($1ft$). If there is a risk of blowout along the path (when MDP¿MAP), the alignment is declared not acceptable; a relaxation of this rule is considered if the blowout location is located at shallow depths (less than 10ft / 3m from the ground surface) in the vicinity of the entry and exit points.

The mechanical validation requires calculating the stresses exerted on the product pipe as it is being pulled into the borehole. Similar to the blowout estimation, the pullback stresses are quantified every 0.3m including the contributions of tension and bending.

Buckling is also checked. If the tested alignment (ant) is viable, it becomes a solution and is temporarily stored in the ledger.

After k new solutions have been found, the solution ledger has a total of (n+k) solutions, n from the previous step plus the newly generated solutions in the present generation. The (n+k) solutions are sorted according to their score, and the k solutions with the lowest scores discarded, therefore, at the end of each generation, there are exactly n solutions in the ledger.

This procedure is repeated for each generation, until the convergence criterion has been met or after a fixed number of generations have been computed. At the end of the optimization stage, the solution ledger has the best n viable combinations of parameters found by the algorithm.

8.6 Algorithm evaluation

In order to assess the performance of the algorithm, we study two design scenarios: one in which the entry and exit points are fixed (scenario 1) and one in which the curvature of the central portion of the alignment is fixed (scenario 2). For each scenario, we present two tests: one using synthetic data, for which the optimal solution is known, and one using real data from actual constructed projects, designed without ACO. Project datasets were provided by Haley Aldrich Inc.

In the simulations with synthetic data, the goal is to study the trends of the solutions returned by the algorithm over large domains and generation numbers, therefore we set the solution ledger size (n) to 100 and the number of generations (nG) to 50. Conversely, for the cases with real data, where the ranges of values for the geometric parameters are more restricted in nature, we set n equal to 25 and nG equal to 30.

For every instance tested, 5 new solutions are found by the algorithm at every optimization generation (i.e., k=5). In order to show the asymptotic behavior of the solutions, no convergence criterion was set, i.e. the nG generations of solutions were calculated. For all

the instances, the parameter q was set to 0.45, resulting in probabilities of 9% and 1% to choose the 1st (best) and 25th solutions in the ledger, respectively. The parameter epsilon (ϵ), which controls the standard deviation around the design parameters, was set to 1.25.

8.6.1 Scenario 1: Fixed Entry/Exit HDD points

The drilling path in the first scenario, when the entry and exit points of the alignment are fixed, is modelled as an alignment composed by 5 segments, including 3 straight segments (entry and exit tangents, plus the central segment), and 2 vertical curves, one at each side of the central segment (see Figure 8.2). Once the entry and exit points are assigned, the configuration of the vertical curves and the subsequent depth of the drill path are to be designed and optimized.

Proof of concept: synthetic data

We consider a simple horizontal profile with a constant elevation (100 ft), the entry and exit points are set at stations 0 and 2000 respectively. The range of the entry and exit angles are fixed between 10 and 20 degrees and the radius of curvature of the vertical projection of the curve is set to 1000 ft. We released the mechanical constraints (borehole stability and pipe integrity) in order to find the shortest and longest possible paths of the alignment. Then, the goal of the proof of concept simulations is to evaluate if the algorithm can yield close-to-optimal geometric solutions.

The minimum possible length of the drill path corresponds to an alignment with the shortest possible entry and exit tangents (minimum value set at 50 ft) and the smallest possible magnitude of their angles (10 degrees). Conversely, the longest drill path corresponds to entry/exit tangents with the maximum magnitude of their angles (20 degrees) and the minimum length of the central segment (50 ft). Such alignments are shown in Figure 8.4.

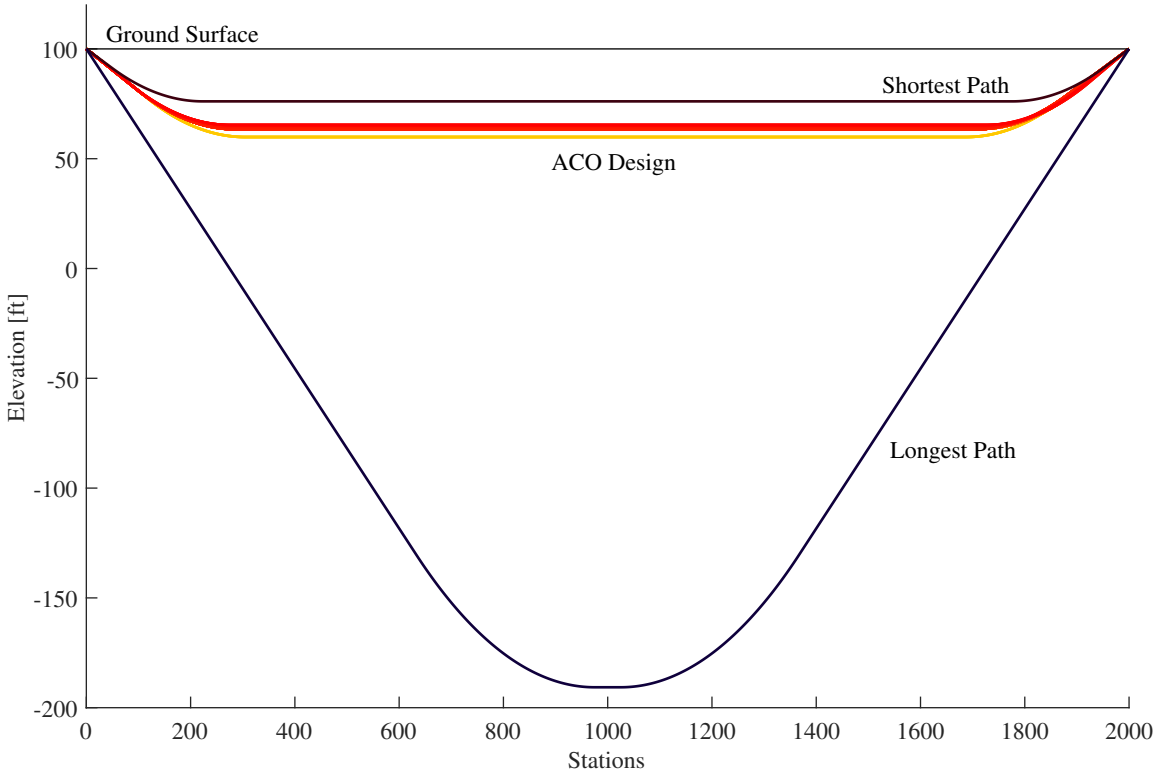

Figure 8.4: Shortest and Longest possible alignments and obtained ACO solutions. Scenario 1: fixed entry/exit points.

In order to show the progressive variation of the variable ranges inside the algorithm, Figure 8.5 shows the evolution of the distribution of the entry angles of the solutions in the ledger as a function of the calculation generations.

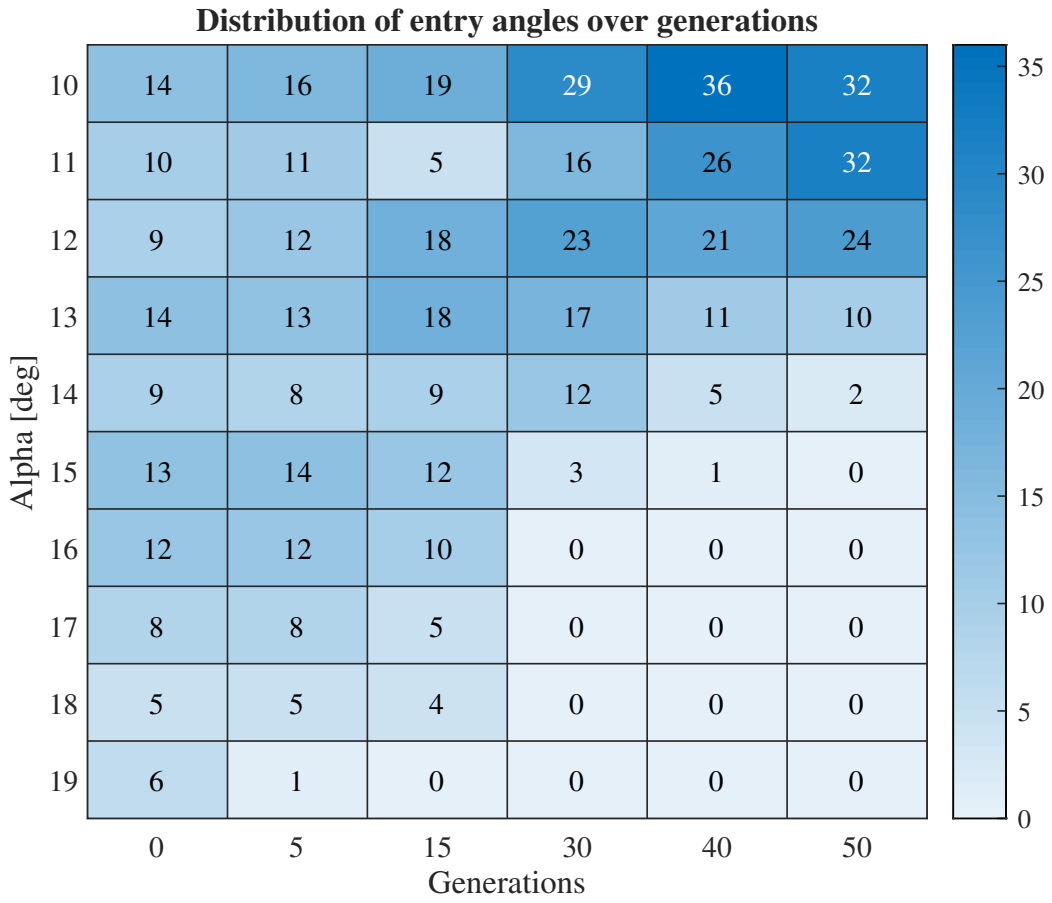

Figure 8.5: Distribution of entry angle over generations. Scenario 1: fixed entry/exit points. Values and color intensity correspond to the percentage of solutions in the corresponding bin, e.g. the sum of values in each column equals to 100.

The entry angles at initialization (generation 0) follow a relatively homogeneous distribution, after which, the distribution progressively narrows down to values closer to 10 degrees, the optimal solution. Lastly, we compare the evolution of the drill length of the ACO solutions against its maximum and minimum values, as shown in Figure 8.6.

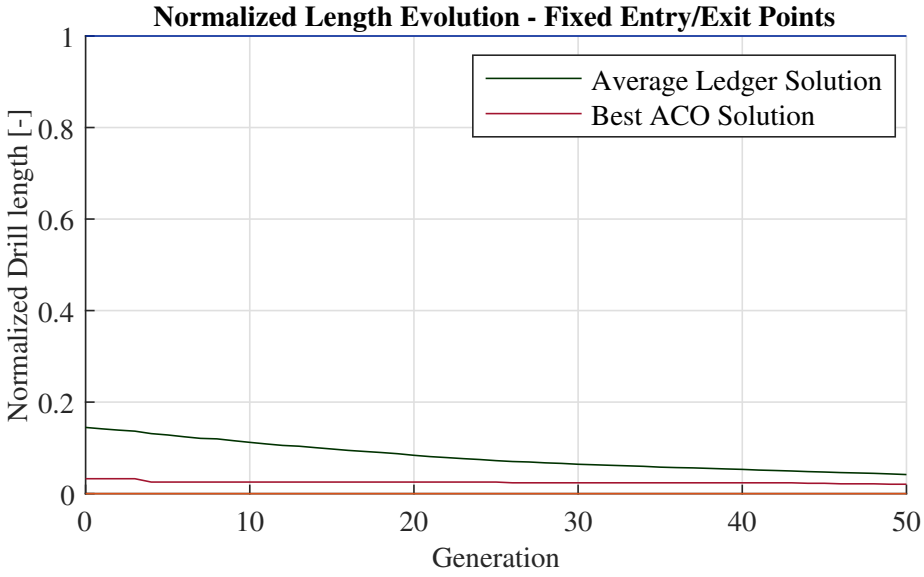

Figure 8.6: Evolution of ACO drill length over generations – Scenario 1: fixed entry/exit points. Normalized drill path length: a value of 1 (respectively 0) means that the drill path length equals the maximum value (respectively, minimum value), at entry/exit angles of 20° (respectively, 10°).

Obtained results show a progressive improvement of the best solution found by ACO resulting in a final drill length that its only 0.1% longer than the global minimum. In the same way, the average length of the solutions in the ledger decreases monotonically, indicating all the solutions in the ledger improve over generations.

8.6.2 Application to real project data

The project considered as an example is a drill passing below a traffic crossing, constructed to connect two sections of an underground power line. The rest of the alignment is built using conventional open trench techniques. The HDD drill hosts a FPVC casing pipe (10 in. in diameter), through which a bundle of power lines will be installed. The total (fixed) plan length of the drill is 2,250 ft (685.8 m). Due to the available equipment and space constraints at the exit site, including pullback area and existing overhead utilities, the range of the entry and exit angles was fixed between 10 and 16 degrees. A minimum depth of cover of 15 ft (4.57 m) and a vertical curve radius of 1,000 ft (304.8m) were adopted for

the alignment.

The properties and characteristics of the product pipe to be installed as well as the equipment and drilling mud are fixed at the first step of the design. Therefore, the optimization step focuses on finding the geometrical configuration of the alignment in terms of the entry/exit angles (α, β) and the bottom elevation of the path (H). Figure 8.7 shows the actual drill path originally designed for the project and the region showing the evolution of the ACO optimized design over the generations (going from lighter to darker colors).

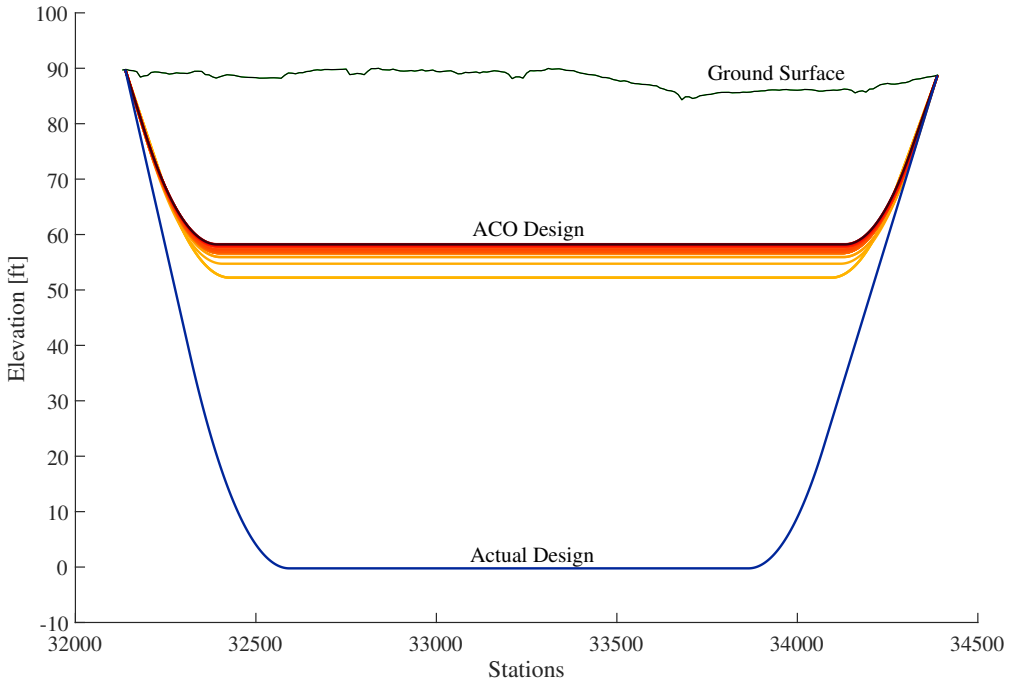

Figure 8.7: Actual and ACO drill path designs. Project in scenario 1: road crossing with fixed entry/exit points. Color intensity on the ACO design illustrates the different generations, going from lighter to darker colors.

Table 8.1 compares the total drill path length and the design parameters found by ACO with the actual design (AD) of the project.

Table 8.1: Scenario 1, summary of design parameters, actual design versus ACO path.

Case	Drill Length [ft]	Bottom Elevation [ft]	Entry Angle [deg]	Exit Angle [deg]
AD	2269.45	-0.23	-16	12
ACO	2254.73	57.82	-11.74	10.76

Obtained results show that even though the actual and the optimized drill paths are considerably different from each other, especially in terms of bottom elevation, the difference in drill path length is relatively small (under 15ft / 4.76m). This result shows that due to the fixed entry and exit points, this project offers little room for minimization of the drill path length.

In order to analyze the performance of the algorithm over the optimization generations, Figure 8.8 shows the evolution of the minimum drill path calculated by the algorithm against the actual drill path originally designed for the project. The geometric minimum corresponds to the shortest possible drill path to connect the entry and exit points and is shown only as a lower bound for reference since it is not a mechanically viable alignment.

In order to capture the distribution of the drill length of the solutions in the ledger, we study the evolution of the interquartile range (IQR), which corresponds to the difference between the values of the 3rd and 1st quartiles of the data, normalized by the geometric minimum bound. Large IQR values indicate that the algorithm has not converged yet, since significantly different drill lengths in the ledger imply a large variability of the design parameters being optimized. Similarly, the mean-min difference is another measure of dispersion, this time quantifying the difference between the average and the minimum drill path lengths among the ledger at a given generation; large values suggest that further refinement of the solution domain is possible. The behavior of the dispersion and variability of the solutions in the ledger is a better indicator of convergence than the evolution of the best design (minimum drill length) alone since it captures the statistically representative state of the algorithm evolution, not a single solution.

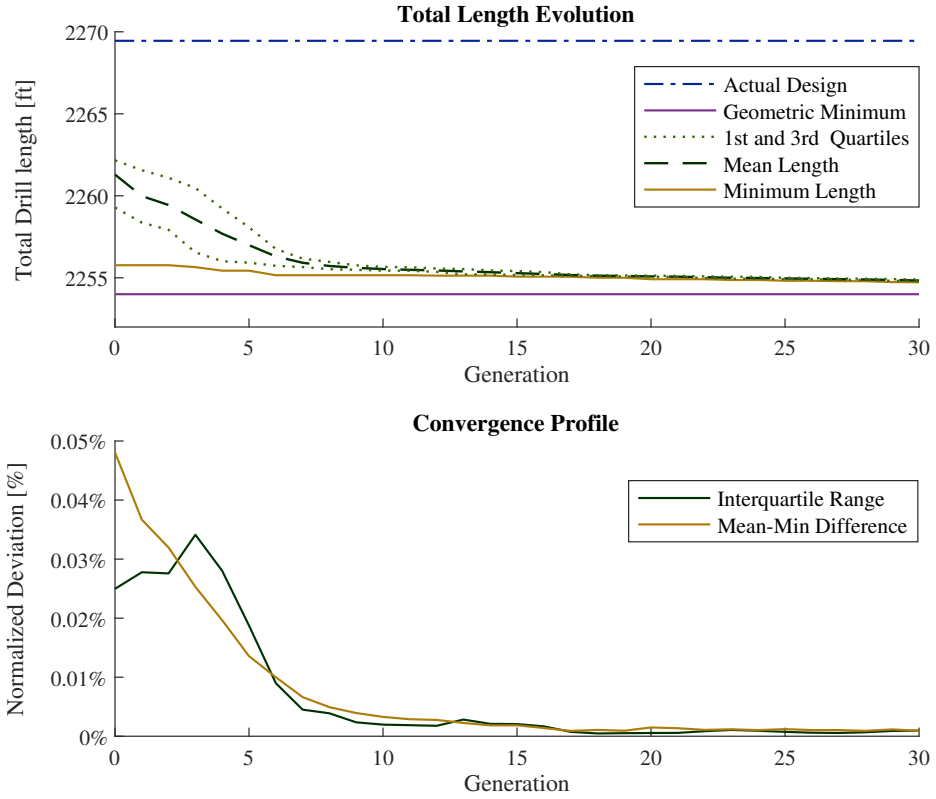

Figure 8.8: Drill path length evolution and convergence. Project in scenario 1: traffic crossing with fixed entry/exit points. Internal dispersion between the solution ledger and minimum drill path length reduce and stabilize after 11 optimization generations.

Figure 8.8 shows that the minimum drill path length found by ACO stabilizes after 11 generations, remaining relatively constant afterwards; due to the small room for minimization of this specific example, the obtained reduction in drill length over generations is about 5 ft (1.5m). The convergence profile shows low dispersion among the solutions in the ledger starting at the first generation. The dispersion is further reduced, to reach a stable optimized path with low dispersion.

The obtained optimized drill path is shorter than the originally designed path, while exhibiting a consistent convergence profile, showing the capability of the algorithm to automate the design process and reach convergence and minimization of the drill path even under highly restricted domains. Results from Figure 8.9, which illustrates the evolution

of the best combination of design parameters, such that they yield the minimum drill path length, show relatively constant values of entry and exit angles (α and β) while the bottom path elevation of the drill path (controlled by H) significantly changes over generations, even after the total drill path appears to stabilize.

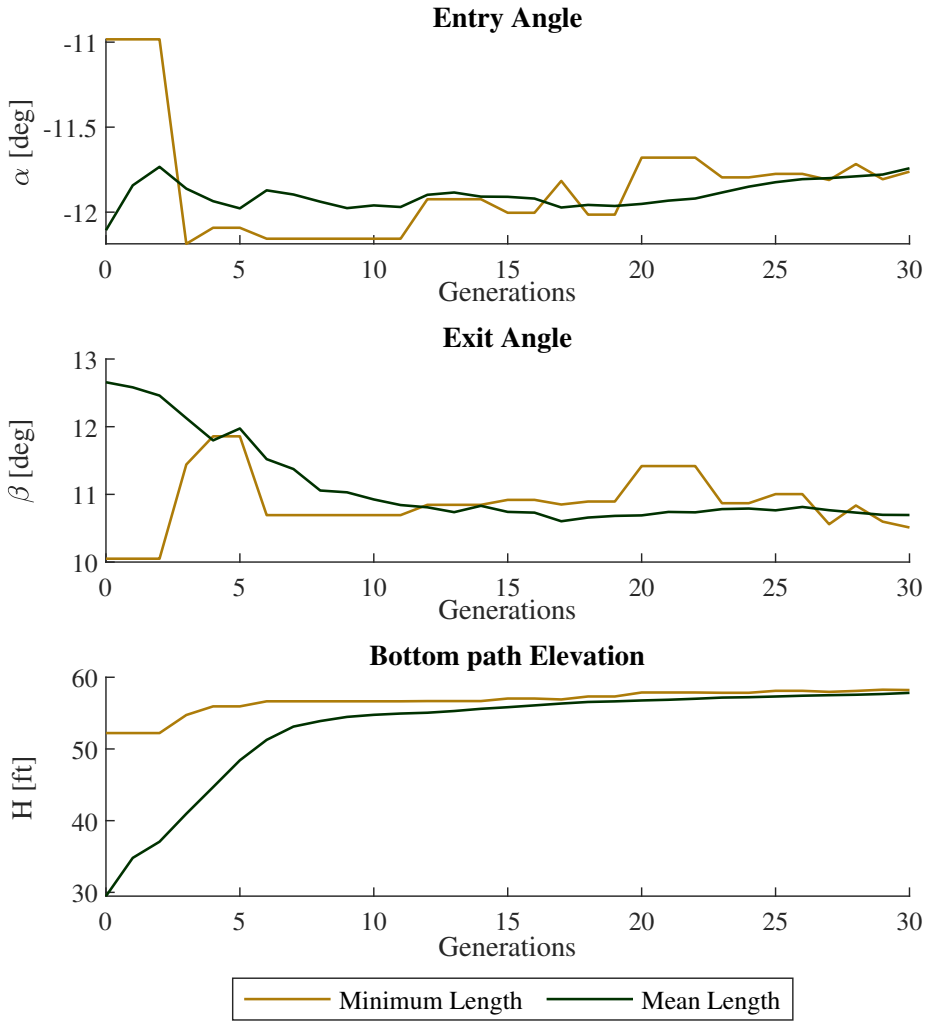

Figure 8.9: Evolution of optimized design parameters. Project in scenario 1: traffic crossing with fixed entry/exit points. Results show large variations of H while α and β show a small range of fluctuation.

8.6.3 Scenario 2: Constrained geometry of central segments

Here, the location of the entry and exit points can be adjusted within a certain interval. This design scenario is particularly useful when considering the existence of physical obstacles

that the HDD alignment must avoid, or designs where the entry and exit azimuth of the alignment are different, in which case, horizontal or compound curves are necessary, and oftentimes located in the central region. Constraints on the entry and exit points depend on the accessibility conditions on site and are transferred to the algorithm in the form of two intervals for the allowable locations of the entry and exit points.

Proof of concept: synthetic data

We consider the same surface profile as the one described in subsection 8.6.2 and an alignment with a single, horizontal segment in the intermediate region spanning from station 750 to 1250 (500 ft in length). The range of the entry and exit angles was fixed between 10 and 20 degrees and the radius of the vertical curve was set to 1000 ft. The entry/exit points of the drill can be placed within 375 ft from the start/end of the alignment respectively. Once again, we released the mechanical constraints (borehole stability and pipe integrity) and compared the results obtained by the algorithm to the global minimum and maximum drill path lengths. The shortest and longest drilling alignments, in addition to the best alignment obtained by the ACO algorithm are shown in Figure 8.10.

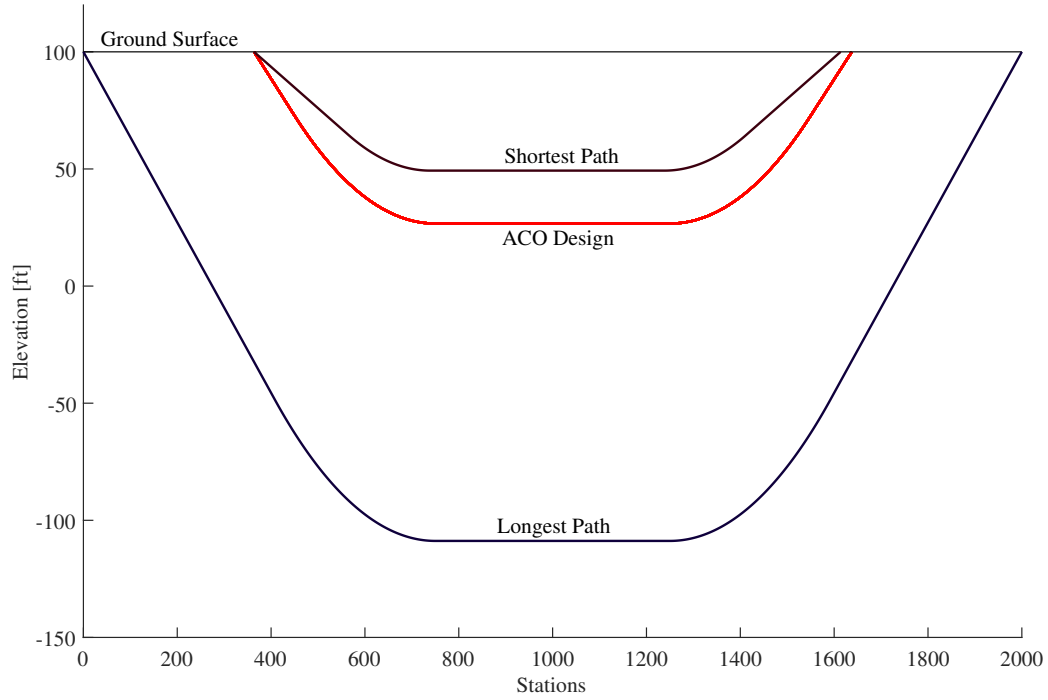

Figure 8.10: Shortest and Longest possible alignments and obtained ACO solutions. Scenario 2: Constrained geometry of the central segment.

Next, we show the evolution of the distribution of the vertical angle of the entry segment along the 50 calculation generations. The results are shown in Figure 8.11.

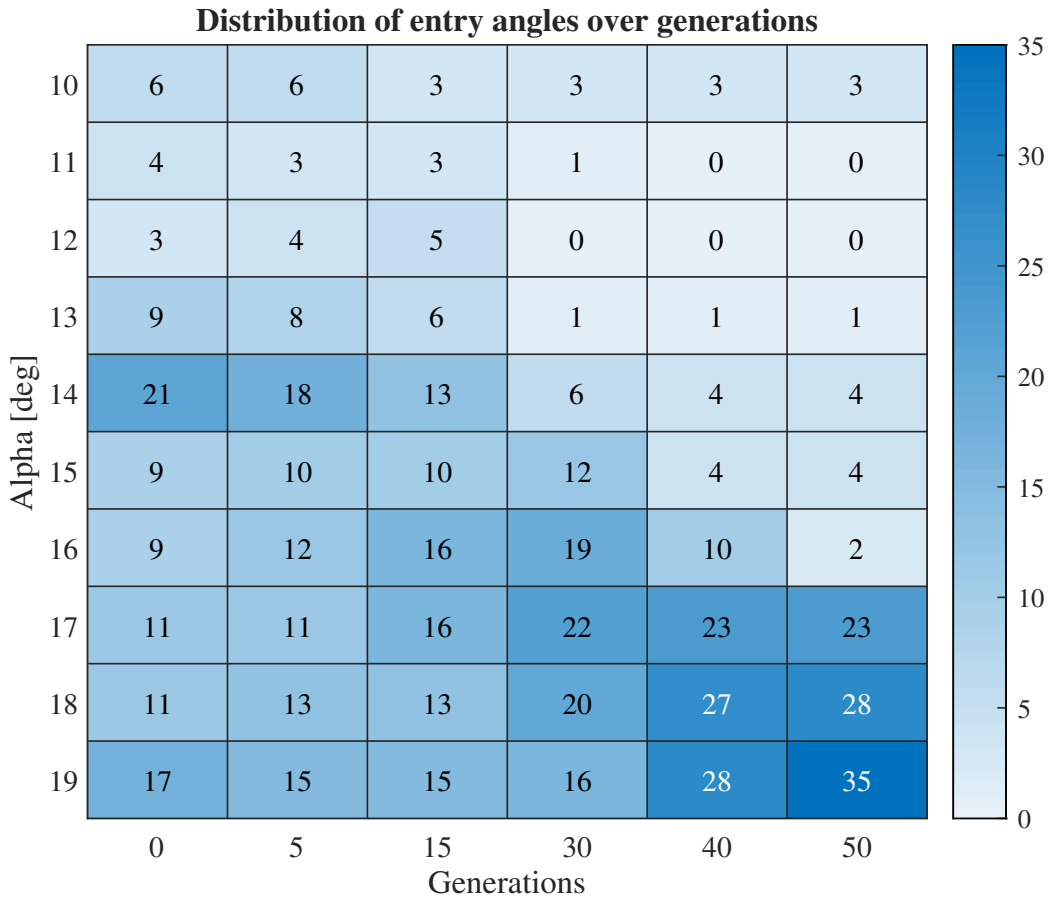

Figure 8.11: Distribution of entry angle over generations. Scenario 2: constrained geometry of the central segment. Values and color intensity correspond to the percentage of solutions in the corresponding bin, e.g. the cum of values in each column equals to 100.

Figure 8.11 shows a relatively uniform distribution of values at initialization, which shifts into a narrow distribution at later generations, suggesting the algorithm has centered around a small range of values. Then, we compare the total length of the solutions found by the algorithm over the generations to the minimum and maximum lengths; the best and average drill path lengths are shown in Figure 8.12.

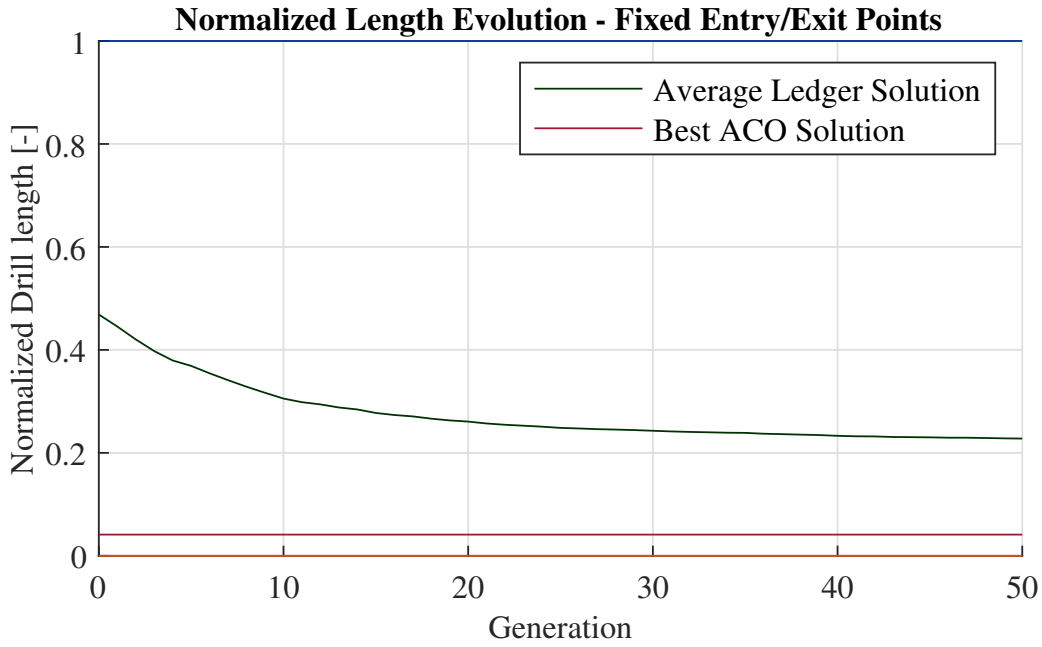

Figure 8.12: Evolution of ACO drill length over generations – Scenario 2: constrained geometry of the central segment. Normalized drill path length: a value of 1 (respectively 0) means that the drill path length equals the maximum value (respectively, minimum value), at entry/exit angles of 20° (respectively, 10°).

Figure 8.12 shows that the average length of the solutions in the ledger decreases as the number of generations increases. However, one of the initial random solutions generated proved to be the best, even after running the different generations, such solution had a total drill length that is only 2.65% longer than the optimal minimum. It is worth noting that this second scenario, with a flexible location of the entry/exit points, yields a significantly higher variability of the solution domain compared to scenario 1, with fixed entry/exit points.

Application to real project data

The project taken as an example here is a drill passing below a shipping channel in an industrial port area, constructed to connect a power line from an electric substation to the rest of the underground alignment. The rest of the project is built using conventional open trench techniques parallel to an existing road. The presence of several foundations from

surrounding structures around the shipping canal constrains the design of the drill alignment. Therefore, a fixed plane central region of the alignment includes a compound helix curve with an 1800 ft (568.4 m) radius, a vertical slope of 0.7 degrees and a 16 degrees horizontal angle. Two adjacent straight segments with the same slope and lengths of 23.9 ft (7.3 m) and 13.48 ft (4.1 m) at the entry and exits point sides complete the central region, Figure 2 shows the configuration of such an alignment in plan and profile views.

Besides the three segments described in the central region of the alignment, four other segments complement the drill path, two vertical curves and two straight segments, one of each at each side of the central region. The HDD drill hosts a steel casing pipe (34 in in diameter) through which a bundle of power lines will be installed. The range of the entry and exit angles was fixed between 8 and 14 degrees. A minimum depth of cover of 15ft (4.57 m) was determined for the design, and the radius of the entry and exit vertical curves was fixed to 1,800 ft (568.4 m), the same as the horizontal radius of the helix curve.

In order to test the effect of the constraint imposed on the location of the entry and exit points, two different instances of the algorithm were tested, one with a 50 ft. viability range for both the entry and exit points (ACO – 50 ft.), and the other with a 150 ft. range for both the entry and exit points (ACO – 150 ft,). Table 8.2 shows the allowable limits for the location of the entry and exit points for each instance.

Table 8.2: Scenario 2: Entry and exit drill points viability ranges.

Case	**Entry Area**			**Exit Area**		
	Min	**Max**	**Actual**	**Min**	**Max**	**Actual**
AD	-	-	50	-	-	1942.7
ACO – 50ft	50	100	99.8	1900	1950	1904.8
ACO – 150ft	50	200	184.0	1800	1950	1802.3

Figure 8.13 shows the actual drill path originally designed for the project, and the evolution of the ACO optimized design over the generations, for each of the two instances. Table 8.3 compares the total drill path length and the design parameters of the actual de-

sign (AD) with those of each of the two ACO instances.

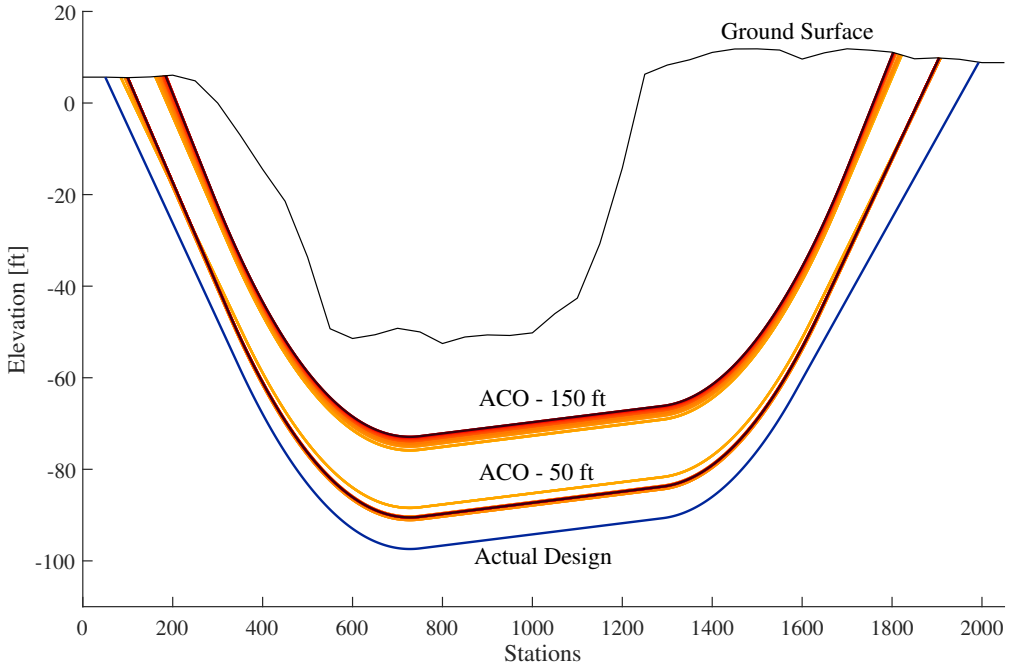

Figure 8.13: Actual and ACO drill path designs. Project in Scenario 2: river crossing with fixed central region design and flexible entry/exit points. Results shown for two ACO instances with different flexibility ranges. Color intensity on the ACO designs corresponds to subsequent generations, going from lighter to darker colors.

Table 8.3: Scenario 2, summary of design parameters, actual design (AD) versus ACO instances.

Case	**Plane Length [ft]**	**Drill Length [ft]**	**Bottom Elevation [ft]**	**Entry Angle [deg]**	**Exit Angle [deg]**
AD	1892.7	1960	-97.27	-12	10
ACO – 50ft	1805	1822.8	-90.27	-12.78	11.75
ACO – 150ft	1618.3	1634.3	-72.89	-13.75	13.76

Obtained optimization results show that setting the drill path length as the optimization objective results in the minimization of the plane length as well, by selecting entry and exit locations as close as possible to the boundaries of the viable ranges. Not surprisingly, a higher flexibility results in reduced drill path length and as the variability range increases,

the optimized design parameters increasingly differ from the actual design (in this case, steeper entry/exit tangents and a significantly shallower drill are obtained by ACO design). Figure 8.14 shows the distribution of the drill path length stored in the solutions ledger over the generations. In a similar way to Scenario 1, the evolution of the minimum (best) length, mean and the 1st and 3rd quartile values are shown over the 30 optimization generations. Given that design ranges for the entry/exit points locations are different for the different instances, their results are not directly comparable to each other, and are meant to illustrate the effect of different design considerations on the performance of the algorithm.

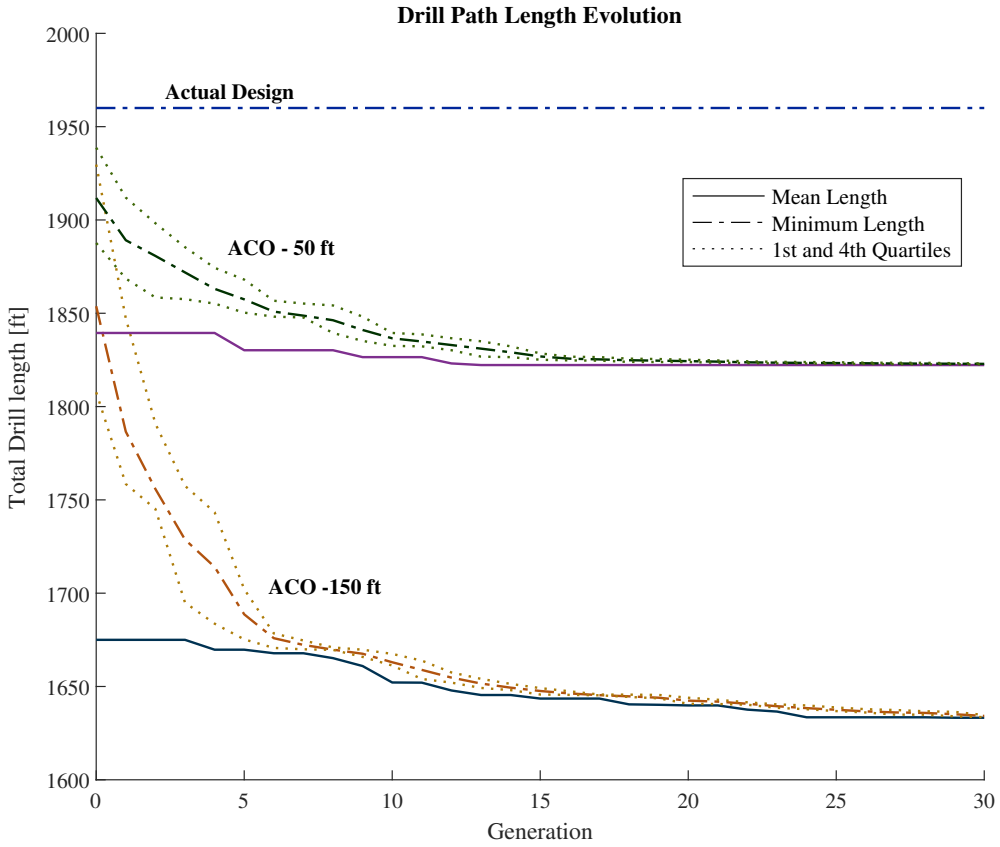

Figure 8.14: Drill path length evolution. Project in Scenario 2: river crossing with two different instances with different entry/exit point location flexibility ranges. Increased flexibility results in shorter paths and increased room for parameter optimization, at the expense of slower convergence.

Results from Figure 8.14 show the influence of the location of the entry/exit points in the drill path length. As expected, instances with shorter allowable plane lengths yield

shorter drill path lengths. Furthermore, the broader search domain created by the increased flexibility of the entry and exit points results in greater minimization of the initial mean drill path length over generations. In terms of convergence, the minimum optimized drill path length stabilized at generation 12 for ACO -50 ft, while it stabilized on generation 23 for ACO – 150 ft. Figure 8.15 shows the evolution of the dispersion measures (IQR and mean-min difference) but this time normalized by the minimum drill length for each instance.

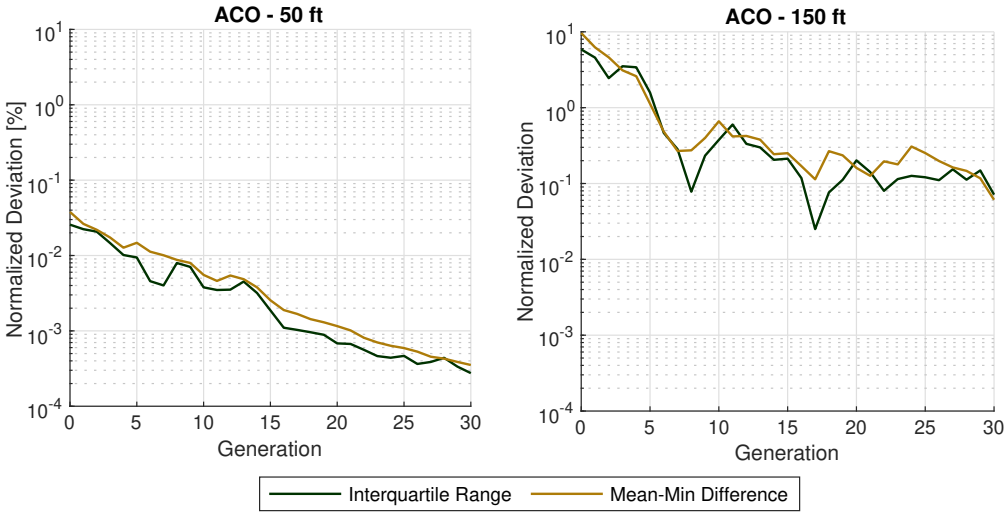

Figure 8.15: Evolution of path length dispersion indexes. Project in Scenario 2: river crossing with two different instances with different entry/exit point location flexibility ranges. Increased flexibility significantly increases the internal solution ledger deviation.

Results from the convergence profile in Figure 8.15 show a significantly larger dispersion among the ledger solutions of ACO – 150 ft. compared to ACO – 50 ft. In both instances though, the dispersion consistently decreases over the generations. This fact suggests that as the flexibility of the domain increases, the potential of minimization increases, enlarging the search domain and potentially improving the optimized solution; but this is at the expense of slower convergence.

This observation is supported by the evolution of the optimal design parameters, as shown in Figure 8.16. The variation of each parameter from the first generation to the last is considerably larger for the more flexible instance (ACO -150 ft) compared to the

more restricted instance (ACO-50 ft). The values of the optimal design parameters seem to stabilize after the first 10 generations for both instances, showing the point at which the exploration of different paths stops, and the algorithm starts converging.

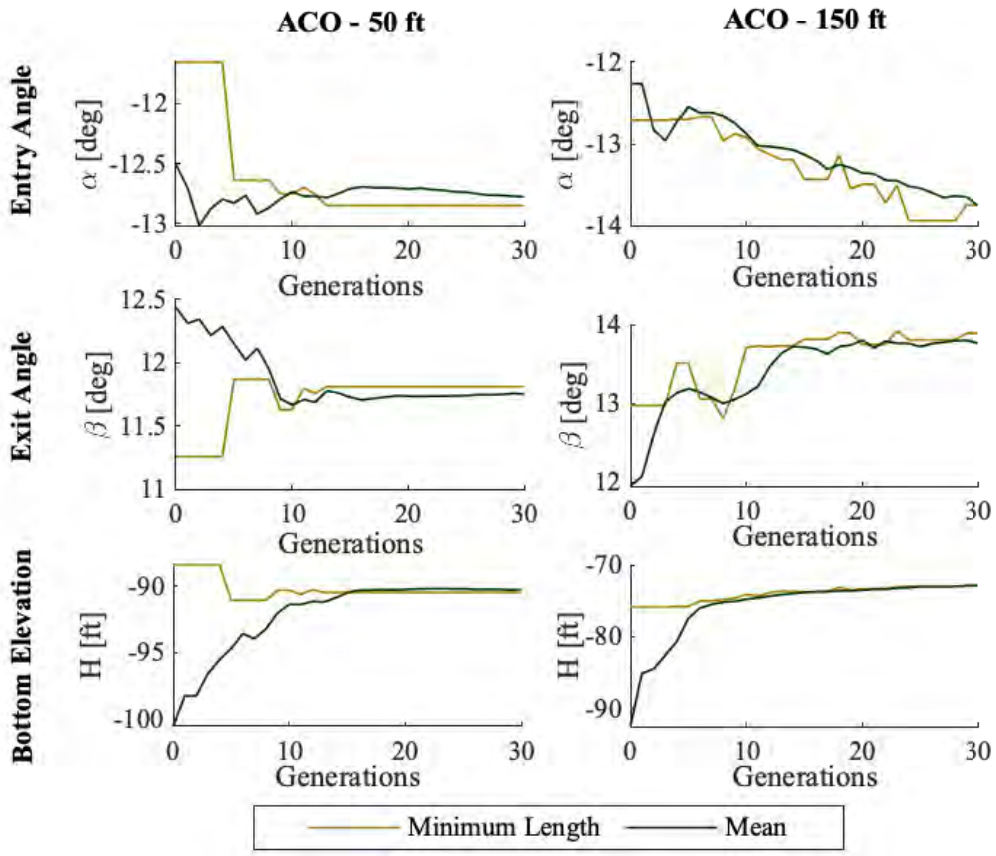

Figure 8.16: Evolution of optimized design parameters. Project in Scenario 2: river crossing with two different instances with different entry/exit point location flexibility ranges. Increased flexibility results in broader fluctuation ranges for optimization parameters.

8.7 Conclusions

Horizontal Directional Drilling (HDD) is gaining increasing interest in the engineering practice because of its relatively low cost, environmental impact, land use and project timeline. Despite its efficiency, HDD relies on empirical and iterative design. To aid routine design procedures, we applied an Ant Colony Optimization (ACO) algorithm to automatically minimize the drill path length under given geometrical constraints (such as

the underground path, that might have to fit between obstacles, or the entry and exit points, which have to be accessible on site), and mechanical constraints (to avoid pipe failure or instability).

Results show the potential of heuristic algorithms like the ACO implementation to automate the design process of HDD alignments and leverage the need of manual iterations from the design engineer. In terms of minimization of the drill path length, the algorithm consistently minimized the drill length over the computation generations, showing a more pronounced decrease in open domains where the design parameters have high flexibility.

The algorithm reached an asymptotic behavior in all the tested instances, showing that broader (less restricted) domains translate into more extensive exploration steps and therefore need more computation generations to reach an asymptotic behavior and convergence. In terms of running time, the algorithm is affected not only by the convergence parameters but also by the ease of finding new valid solutions, which depends on the geometric configuration and input parameters of the design. Nevertheless, the optimization algorithm does not incur in extensive computation times, being always under 15 minutes of run time using a serial implementation on a 2.40 GHz machine (Intel i5-6300U) for each of the instances tested.

A natural extension of the present work will be to release the constraint that imposes that the optimization of the path occurs along a fixed alignment in plane view, which must be decided a priori by the design engineer. Future implementations could include a more comprehensive search domain, which instead of being limited to a 2D (fixed) geotechnical profile, could include a 3D domain to find the best path between two allowable entry and exit areas. Nonetheless, this extension of the algorithm requires a significantly more extensive knowledge of the topographic and geotechnical characteristics of the 3D construction domain.

Lastly, it is important to highlight that the accuracy and viability of the algorithm is based on the set of constraints and variables on which the algorithm is built on, therefore

the current implementation does not intend to replace the judgment of a trained engineer who must validate the input parameters and results obtained from the algorithm.

CHAPTER 9
CONCLUSIONS

The present thesis proposes different methods to optimize underground infrastructure. First, we model infrastructure as a network that can be modeled and designed using different bio-inspired algorithms, depending on the functional goals of the network itself. Second, we observe, interpret and model the deformation and failure of shallow underground cavities, and we apply machine learning and bio-inspired algorithms to optimize their design. In the following, we offer some concluding remarks on each of the thesis chapters, on the current limitations, and on what we envision as future work on the topic.

9.1 Part 1: Bio-inspired optimization of transport networks

The first part of the thesis assesses the feasibility of applying biological principles to improve the design of 'human-made' underground networks, and, more generally, transportation systems. To this end, we cover the different steps of bio inspiration, first, studying the fundamental behavior of a type of natural network, slime mold, then characterizing its behavior as a transportation network, and lastly, exploring two different applications of such biological network dynamics to engineering problems: the extraction of resources from a reservoir, and the deployment of transport networks in a city.

The study of slime mold morphogenesis in attractive, adverse and neutral environments in chapter 2 shows that the substrate composition directs the behavior of slime molds. Slime molds in neutral environments proved to encourage exploration of the domain, while in attractive substrates (nutritive), the cell prioritizes metabolization of the nutrient present in the substrate, which significantly slows down the spreading rate. On the other hand, the presence of an adverse substrate increases the pulsatile behavior of the cell, which alternated cycles of growth and retraction, with a low net spreading rate overall, and a

limited exposure to the adverse chemical present on the substrate. The spreading direction of slime mold is initially random, but, depending on the substrate, the mucus left behind by the cell acts as a spatial memory, influencing the behavior of the cell in subsequent steps of growth. For instance, in neutral environments, the exploration of new areas is preferred to zones already explored (mucus); while in adverse environments, the growth over a mucus layer is preferred since it provides a layer of protection that minimizes the exposure to the repellent distributed over the substrate. Lastly, experiments with discrete sets of attractive and adverse spots over a neutral substrate did not affect the overall behavior of the cell, suggesting that the behavior is controlled by the substrate, not the presence of localized obstacles or attraction points.

After understanding the fundamental behavior behind slime mold expansion, the next step is to characterize its topology as a network. chapter 3 presents a study such networks in neutral, adverse and nutritive environments, as well as an analysis of the process of fusion between independent slime mold cells. The findings from this characterization further validate the observations from chapter 2, showing that cells in a neutral environment are more 'spread-out', developing networks that cover a significantly larger area than the area of slime mold itself, with thin and long veins and an overall higher network length. On the contrary, cells in adverse and nutritive substrates are denser and more compact, have thicker veins and are more centralized. The connectivity of the networks, in terms of node degree and alpha coefficient, is similar across substrates. The total network length is close to that of the minimum spanning tree, showing that even though the network expands and retracts in cycles, the total wiring cost is kept close to a minimum. In terms of efficiency, the networks in neutral and adverse environments performed better than the ones in nutritive conditions, while the resiliency of the networks proved to be correlated with the total network length, showing that highly interconnected networks are able to sustain more damage before becoming disconnected. The results on fusion show that after the initial contact, the cells fuse, and start refining, with significant differences between substrates. Cells in

neutral environments refine the fusion region to keep several thin veins connecting both cells, while adverse environments minimize the number, keeping only 1 or 2. The number of connecting veins is also reflected in the path efficiency between cells, which is better in neutral, well connected networks.

These results show that even though slime molds are relatively simple organisms, they are capable to build complex, dynamic and efficient networks that adapt to different environments while remaining resilient and efficient. In addition, such networks refine their topology based on fluctuating environmental conditions. This behavior could inspire the design of infrastructure transportation networks, which not only follow similar optimization goals, e.g. transport efficiency, cost reduction and resiliency, but are also embedded in environments that fluctuate. The network adaptability exhibited by slime mold is desirable in civil engineering. The dynamics of slime mold networks, and their fusion in response to various environmental constraints has a potential to inspire new strategies to design adaptable information and infrastructure networks, that are resilient to natural and biological hazards as well as geopolitical risks. One slime mold inspired application is presented in chapter 4. An algorithm that models the feedback mechanism of slime mold (by regulating refinement and strengthening of network veins) was applied to optimize the extraction (or injection) of resources in porous media. The results show that the algorithm successfully creates the optimal path to extract the resources from the porous medium, reaching up to 32% of the distributed resource by using only 10% of the edges in the network. In addition, the algorithm works successfully in heterogeneous domains, creating direct and wide flow paths that minimize drag through low resource, low conductivity regions, while creating branching, dendritic paths in regions with high resource concentration, that maximize its extraction efficiency. Future work on this topic can focus on adding constraints to the developed solution, in order to account for constructive, economic and environmental factors.

chapter 5 shows a different application of bio-inspired networks, this time using a leaf

venation (LV) growth algorithm to design a transportation network that would serve equitably the population of the city of Atlanta, and to extend the actual rail lines to an adjacent county. Results were compared to global and local optimization algorithms, showing that the leaf venation network is almost Pareto optimal, achieving an almost-optimal trade-off between transport efficiency and network cost. Moreover, the LV networks outperform the global optimal network (Steiner tree), specially when networks are highly centralized around a node, equivalent to the petiole in leaves. When considering the total network cost as a function of length and traffic through the edges, LVs yield networks that are less expensive than Steiner trees, and the difference is accentuated as networks are more centralized around a certain node, which in this case corresponds to the closest rail station of the existing system. Future work on this topic could consider extra variables, for instance, the cost of land, topography, matrix origin/destination, commuting trends, and demographic projections.

9.2 Part 2: Mechanics of shallow underground cavities

The second part of the thesis focuses on the mechanical response of shallow underground cavities. The main conclusions of this thesis call for a new way of designing geotechnical infrastructure, without forfeiting a rigorous mechanical understanding of the processes.

First, chapter 6 shows an experimental study on the deformation and failure mechanisms of granular soil around pressurized shallow cavities, using x-ray computed tomography. Results demonstrate that the failure mechanism of shallow cavities does not follow the commonly assumed cavity expansion failure mechanism in an infinite domain. Instead, it exhibits shear bands that follow a catenary shape and reach the free surface. The soil inside such bands either dilates or contracts, depending on the initial density of the material. Similarly, the orientation of strain increments follows catenary trajectories along the cross-section of the cavity, so long as plane strain conditions are preserved. The extent of the cavity length under plane strain conditions was at most 75% of the total cavity length, and

turned out to be a function of cavity length, soil density, and vertical load. Results could inform further analytical developments, which oftentimes rely on plane-strain assumptions, knowledge of the shape of the failure planes, and co-axiality of strain and stress. A possible application is the prediction of the blowout pressure of shallow cavities such as horizontal directional drills.

The increased mathematical complexity of the specific loading conditions limits the availability of accurate (pseudo-) analytical methods to predict stress and strain fields around cavities, and calls for complementary numerical techniques. Chapter chapter 7 further studies the behavior of shallow cylindrical cavities, this time using the finite element method coupled with machine learning algorithms. The results shown in this study exemplify the benefits of using machine learning algorithms to generate prediction tools that can better inform design decisions. In particular, a trained model could be used to predict whether a given cavity is prone to blowout, classifying its susceptibility into three categories: low, mid and high. When the cavity falls in the low (respectively, high) category, the design is validated (respectively, invalidated). The mid category calls for a more in-depth assessment, for instance, by creating a dedicated FEM model to check the mechanical response of the designed cavity. Machine learning algorithms to readily assess underground cavity stability offer a new toolbox of methods towards safer, more efficient design. These new methods are not intended to replace rigorous mechanical analyses, but can be used to bring attention to scenarios in which these strict analyses are needed, providing an extra layer of safety to the design process. Possible future work includes, and is not limited to, similar analyses oriented towards in-situ soil characterization, stability assessment considering the stochastic distribution of soil properties and autonomous systems that can adjust their burrowing/drilling mechanisms on the fly, without the need of constant feedback from an operator.

The last chapter of this thesis (chapter 8) showcases the industry application of an automated optimization algorithm. The bio-inspired algorithm implemented (ant colony opti-

mization) is used to automate the design process of a horizontal directional drill, checking the mechanical stability of the cavities and the construction constraints, while at the same time, minimizing the total cost of the tunnel. The algorithm was implemented and tested in partnership with Haley Aldrich Inc, using data from two real projects. The results of the validation of the method showed three main benefits of the algorithm in comparison to the traditional design methods. First, the obtained alignment design was at least 10% shorter than the one obtained without the algorithm, which yields cost savings. Second, the method increases the accuracy of the design. Third, the design time is decreased to at least a third, depending of the case benchmarked, leveraging the need of manual labor and shifting the use of the design engineer to checking and validating the results, instead of focusing on repetitive calculation tasks. It is important to highlight that the accuracy of the algorithm is based on the set of constraints and variables assigned to it, and therefore, the current implementation should not replace the judgment of a trained engineer who must validate the results. A natural extension of the present work could release the geometric constraints, currently limited to a two-dimensional profile, to a three-dimensional domain. Another possible perspective is the integration of the different steps of the process, using GIS, CAD and BIM tools, in order to develop a more complete and robust design tool.

The overarching conclusion of this thesis is the feasibility and applicability of bio-inspiration towards the optimization and improvement of our current design techniques in geotechnical engineering. Moreover, the results call for an integration of conception and execution of projects. The biggest gap to bridge in the design of underground cavities, and infrastructure in general, is the system-level conception, and the integration of knowledge from biology and other disciplines.

REFERENCES

[1] D. Kaliampakos, A. Benardos, and A. Mavrikos, "A review on the economics of underground space utilization," *Tunnelling and Underground Space Technology*, vol. 55, pp. 236–244, May 2016.

[2] R. Guidotti, H. Chmielewski, V. Unnikrishnan, P. Gardoni, T. McAllister, and J. van de Lindt, "Modeling the resilience of critical infrastructure: The role of network dependencies," *Sustainable and resilient infrastructure*, vol. 1, no. 3-4, pp. 153–168, 2016.

[3] W. Broere, "Urban underground space: Solving the problems of today's cities," *Tunnelling and Underground Space Technology*, vol. 55, pp. 245–248, May 2016.

[4] Y.-K. Qiao, F.-L. Peng, and Y. Wang, "Monetary valuation of urban underground space: A critical issue for the decision-making of urban underground space development," *Land Use Policy*, vol. 69, pp. 12–24, Dec. 2017.

[5] J. Peng and F. L. Peng, "A GIS-based evaluation method of underground space resources for urban spatial planning: Part 1 methodology," *Tunnelling and Underground Space Technology*, vol. 74, pp. 82–95, Apr. 2018.

[6] Y. Bar-Cohen, *Biomimetics: nature-based innovation*. CRC press, 2011.

[7] S. Hawking, *The illustrated a brief history of time*. Bantam, 1996.

[8] G. M. Whitesides, "Bioinspiration: something for everyone," *Interface Focus*, vol. 5, no. 4, p. 20 150 031, Aug. 2015.

[9] C. Edvardsen, "Water permeability and autogenous healing of cracks in concrete," in *Innovation in concrete structures: Design and construction*, Thomas Telford Publishing, 1999, pp. 473–487.

[10] "A bio-inspired perspective for geotechnical engineering innovation," in *Geotechnical Special Publication*, American Society of Civil Engineers (ASCE), 2017, pp. 862–870.

[11] R. Vepa, "Biomimetic flight and flow control: Learning from the birds," in *Solid Mechanics and its Applications*, vol. 7, Springer, Dordrecht, 2008, pp. 443–447, ISBN: 9781402068577.

[12] N. F. Lepora, P. Verschure, and T. J. Prescott, "The state of the art in biomimetics," *Bioinspiration & Biomimetics*, vol. 8, no. 1, p. 013 001, Jan. 2013.

[13] D. S. Halacy, *Bionics, the Science of" living" Machines*. Holiday House, 1965.

[14] Y. Bar-Cohen, "Biomimetics—using nature to inspire human innovation," *Bioinspiration & Biomimetics*, vol. 1, no. 1, P1–P12, Mar. 2006.

[15] R. Allen, "From feathers to fins: can we understand biological systems—and learn from them?" *Bioinspiration & Biomimetics*, vol. 1, no. 1, E01, Mar. 2006.

[16] J. T. DeJong and E. Kavazanjian, "Bio-mediated and Bio-inspired Geotechnics," in, Springer, Cham, 2019, pp. 193–207.

[17] S. D. Mallett, S. Matsumura, and J. David Frost, "Additive manufacturing and computed tomography of bio-inspired anchorage systems," *Geotechnique Letters*, vol. 8, no. 3, pp. 219–225, 2018.

[18] A. A. QABANY and K. Soga, "Effect of chemical treatment used in micp on engineering properties of cemented soils," in *Bio-and Chemo-Mechanical Processes in Geotechnical Engineering: Géotechnique Symposium in Print 2013*, ICE Publishing, 2014, pp. 107–115.

[19] M. Burrall, J. T. DeJong, A. Martinez, D. W. Wilson, and L. Huang, "Bio-inspiration through tree root pullout tests for innovative anchorage design," in *Geo-Congress 2020: Biogeotechnics*, American Society of Civil Engineers Reston, VA, 2020, pp. 233–242.

[20] L. Huang and A. Martinez, "Study of interface frictional anisotropy at bioinspired soil-structure interfaces with compliant asperities," in *Geo-Congress 2020: Biogeotechnics*, American Society of Civil Engineers Reston, VA, 2020, pp. 253–261.

[21] A. Martinez, K. O'Hara, S. Sinha, D. Wilson, and K. Ziotopoulou, "Monotonic and cyclic centrifuge testing of snake skin-inspired piles," in *Proc., Biomed. and Bioinsp. Geotech.(B2G) Conf*, 2018.

[22] J. D. Frost, M. M. Roozbahani, K. Jackson, L. Leonard, K. Yamamoto, M. Jones, and A. Martinez, "Biologically-inspired insight into soil arching and tunnel stability from the topology of ant nests," in *press for 19 th Int. Conf. Soil Mech. Geotech. Eng*, 2017.

[23] Y. Ma, T. M. Evans, and D. D. Cortes, "2d dem analysis of the interactions between bio-inspired geo-probe and soil during inflation–deflation cycles," *Granular Matter*, vol. 22, no. 1, p. 11, 2020.

[24] S. Huang and J. Tao, "Modeling clam-inspired burrowing in dry sand using cavity expansion theory and dem," *Acta Geotechnica*, pp. 1–22, 2020.

[25] A. Tero, S. Takagi, T. Saigusa, K. Ito, D. P. Bebber, M. D. Fricker, K. Yumiki, R. Kobayashi, and T. Nakagaki, "Rules for biologically inspired adaptive network design.," *Science (New York, N.Y.)*, vol. 327, no. 5964, pp. 439–42, Jan. 22, 2010.

[26] F. Patino-Ramirez and C. Arson, "Transportation networks inspired by leaf venation algorithms," *Bioinspiration & Biomimetics*, vol. 15, no. 3, p. 036 012, Mar. 2020.

[27] F. Patino-Ramirez, A. Boussard, C. Arson, and A. Dussutour, "Substrate composition directs slime molds behavior," *Scientific Reports*, vol. 9, no. 1, pp. 1–14, Dec. 2019.

[28] F. Patino-Ramirez, A. Dussutour, and C. Arson, "Substrate and cell fusion influence on slime mold network dynamics," *Scientific Reports*, Under Review, 2020.

[29] F. Patino-Ramirez, C. Arson, *et al.*, "Bio-inspired fluid extraction model for reservoir rocks," in *51st US Rock Mechanics/Geomechanics Symposium*, American Rock Mechanics Association, 2017.

[30] R. J. Mair and D. M. Wood, *Pressuremeter Testing : Methods and Interpretation.* 1st. Elsevier Science, 1987, p. 169, ISBN: 9781483102252.

[31] L. Li, J. Li, and D. Sun, "Anisotropically elasto-plastic solution to undrained cylindrical cavity expansion in K0-consolidated clay," *Computers and Geotechnics*, vol. 73, pp. 83–90, 2016.

[32] D. M. Wood, "Strain-dependent mduli and pressuremeter tests," *Géotechnique*, vol. 40, no. 3, pp. 509–512, Sep. 1990.

[33] R. Salgado, J. K. Mitchell, and M. Jamiolkowski, "Cavity Expansion and Penetration Resistance in Sand," *Journal of Geotechnical and Geoenvironmental Engineering*, vol. 123, no. 4, pp. 344–354, Apr. 1997.

[34] A. R. Russell and N. Khalili, "Cavity Expansion Theory and the Cone Penetration Test in Unsaturated Sands," in *Unsaturated Soils 2006*, Reston, VA: American Society of Civil Engineers, Mar. 2006, pp. 2546–2557, ISBN: 9780784408025.

[35] H. Zhou, G. Q. Kong, and H. L. Liu, "Pressure-controlled elliptical cavity expansion under anisotropic initial stress: Elastic solution and its application," *Science China Technological Sciences*, vol. 59, no. 7, pp. 1100–1119, Jul. 2016.

[36] B. Keulen, "Maximum allowable pressures during horizontal directional drillings focused on sands," PhD thesis, TU Delft, 2001, pp. 1–120.

[37] H. W. Xia and I. D. Moore, "Estimation of maximum mud pressure in purely cohesive material during directional drilling," *Geomechanics and Geoengineering*, vol. 1, no. 1, pp. 3–11, Mar. 2006.

[38] A. Rostami, Y. Yi, and A. Bayat, "Estimation of Maximum Annular Pressure during HDD in Noncohesive Soils," *International Journal of Geomechanics*, vol. 17, no. 4, p. 06 016 029, Apr. 2017.

[39] K. Terzaghi, R. B. Peck, and G. Mesri, *Soil Mechanics in Engineering Practice*, 3rd. Wiley, Nov. 1996, p. 534, ISBN: 978-0-471-08658-1.

[40] C. H. Evans, "An examination of arching in granular soils," PhD thesis, Massachusetts Institute of Technology, 1984.

[41] H.-j. Tien, "A Literature Study of the Arching Effect," PhD thesis, Massachusetts Institute of Technology, 1996, pp. 1–196.

[42] F. Patino-Ramirez, G. Viggiani, B. Caicedo, C. Arson, and F. Anselmucci, "Deformation and failure mechanisms of granular soil around pressurised shallow cavities," *Geotechnique*, Under Review,

[43] F. Patino-Ramirez, "Machine learning algorithms applied to the blowout susceptibility estimation around pressurized cavities in drained soil," in *IS 2020: 10th International Symposium on Geotechnical Aspects of Underground Construction in Soft Ground*, 2021.

[44] F. Patino-Ramirez, C. Layhee, and C. Arson, "Horizontal directional drilling (HDD) alignment optimization using ant colony optimization," *Tunnelling and Underground Space Technology*, vol. 103, p. 103 450, Sep. 2020.

[45] P. Ball and N. R. Borley, *The self-made tapestry: pattern formation in nature*. Oxford University Press Oxford, 1999, vol. 198.

[46] M. Rietkerk and J. de Koppel, "Regular pattern formation in real ecosystems," *Trends in ecology & evolution*, vol. 23, no. 3, pp. 169–175, 2008.

[47] B. P. Teague, P. Guye, and R. Weiss, "Synthetic morphogenesis," *Cold Spring Harbor perspectives in biology*, vol. 8, no. 9, a023929, 2016.

[48] G. Theraulaz, J. Gautrais, S. Camazine, and J.-L. Deneubourg, "The formation of spatial patterns in social insects: From simple behaviours to complex structures," *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 361, no. 1807, pp. 1263–1282, 2003.

[49] K. J. Chalut and E. K. Paluch, "The actin cortex: A bridge between cell shape and function," *Developmental cell*, vol. 38, no. 6, pp. 571–573, 2016.

[50] M. K. Driscoll, C. McCann, R. Kopace, T. Homan, J. T. Fourkas, C. Parent, and W. Losert, "Cell shape dynamics: From waves to migration," *PLoS computational biology*, vol. 8, no. 3, e1002392, 2012.

[51] B. C. Goodwin, "Unicellular morphogenesis," *Cell Shape: Determinants, Regulation, and Regulatory Role*, pp. 365–391, 1989.

[52] G. Salbreux, G. Charras, and E. Paluch, "Actin cortex mechanics and cellular morphogenesis," *Trends in cell biology*, vol. 22, no. 10, pp. 536–545, 2012.

[53] G. Charras and E. Sahai, "Physical influences of the extracellular environment on cell migration," *Nature reviews Molecular cell biology*, vol. 15, no. 12, p. 813, 2014.

[54] S. van Helvert, C. Storm, and P. Friedl, "Mechanoreciprocity in cell migration," *Nature cell biology*, vol. 20, no. 1, p. 8, 2018.

[55] P. Friedl and K. Wolf, "Plasticity of cell migration: A multiscale tuning model," *The Journal of cell biology*, vol. 188, no. 1, pp. 11–19, 2010.

[56] M. L. Gardel, I. C. Schneider, Y. Aratyn-Schaus, and C. M. Waterman, "Mechanical integration of actin and adhesion dynamics in cell migration," *Annual review of cell and developmental biology*, vol. 26, pp. 315–333, 2010.

[57] N. Inagaki and H. Katsuno, "Actin waves: Origin of cell polarization and migration?" *Trends in cell biology*, vol. 27, no. 7, pp. 515–526, 2017.

[58] A. Aman and T. Piotrowski, "Cell migration during morphogenesis," *Developmental biology*, vol. 341, no. 1, pp. 20–33, 2010.

[59] G. Reig, E. Pulgar, and M. L. Concha, "Cell migration: From tissue culture to embryos," *Development*, vol. 141, no. 10, pp. 1999–2013, 2014.

[60] K. M. Yamada and R. Mayor, "Cell dynamics in development, tissue remodelling, and cancer," *Current opinion in cell biology*, vol. 42, p. iv, 2016.

[61] L. E. Hind, W. J. B. Vincent, and A. Huttenlocher, "Leading from the back: The role of the uropod in neutrophil polarization and migration," *Developmental cell*, vol. 38, no. 2, pp. 161–169, 2016.

[62] P. A. Iglesias and P. N. Devreotes, "Navigating through models of chemotaxis," *Current opinion in cell biology*, vol. 20, no. 1, pp. 35–40, 2008.

[63] J. S. King and R. H. Insall, "Chemotaxis: Finding the way forward with dictyostelium," *Trends in cell biology*, vol. 19, no. 10, pp. 523–530, 2009.

[64] T. Mseka, J. R. Bamburg, and L. P. Cramer, "Adf/cofilin family proteins control formation of oriented actin-filament bundles in the cell body to trigger fibroblast polarization," *Journal of cell science*, vol. 120, no. 24, pp. 4332–4344, 2007.

[65] S. H. Larsen, J. Adler, J. J. Gargus, and R. W. Hogg, "Chemomechanical coupling without atp: The source of energy for motility and chemotaxis in bacteria," *Proceedings of the National Academy of Sciences*, vol. 71, no. 4, pp. 1239–1243, 1974.

[66] Y. Nakaoka and K. Iwatsuki, "Hyperpolarization-activated inward current associated with the frequency increase in ciliary beating of paramecium," *Journal of Comparative Physiology A*, vol. 170, no. 6, pp. 723–727, 1992.

[67] J. Saranak and K. W. Foster, "Rhodopsin guides fungal phototaxis," *Nature*, vol. 387, no. 6632, p. 465, 1997.

[68] S. Kuroda, S. Takagi, T. Nakagaki, and T. Ueda, "Allometry in physarum plasmodium during free locomotion: Size versus shape, speed and rhythm," *Journal of experimental biology*, vol. 218, no. 23, pp. 3729–3738, 2015.

[69] T. Nakagaki, H. Yamada, and A. Tóth, "Maze-solving by an amoeboid organism.," *Nature*, vol. 407, no. 6803, p. 470, 2000.

[70] B. Rodiek and M. J. B. Hauser, "Migratory behaviour of physarum polycephalum microplasmodia," *The European Physical Journal Special Topics*, vol. 224, no. 7, pp. 1199–1214, 2015.

[71] D. Vogel, A. Dussutour, and J.-L. Deneubourg, "Symmetry breaking and interclonal behavioural variability in a slime mould," *Biology Letters*, vol. 14, no. 12, p. 20 180 504, 2018.

[72] D. Vogel, S. C. Nicolis, A. Perez-Escudero, V. Nanjundiah, D. J. T. Sumpter, and A. Dussutour, "Phenotypic variability in unicellular organisms: From calcium signalling to social behaviour," *Proceedings of the Royal Society B: Biological Sciences*, vol. 282, no. 1819, p. 20 152 322, 2015.

[73] H. Aldrich, *Cell biology of Physarum and Didymium V1: organisms, nucleus, and cell cycle*. Elsevier, 2012.

[74] C. Oettmeier, K. Brix, and H.-G. Döbereiner, "Physarum polycephalum—a new take on a classic model system," *Journal of Physics D: Applied Physics*, vol. 50, no. 41, p. 413 001, 2017.

[75] R. Kuroda and H. Kuroda, "Relation of cytoplasmic calcium to contractility in physarum polycephalum," *Journal of cell science*, vol. 53, no. 1, pp. 37–48, 1982.

[76] T. Ueda, K. Terayama, K. Kurihara, and Y. Kobatake, "Threshold phenomena in chemoreception and taxis in slime mold physarum polycephalum.," *The Journal of general physiology*, vol. 65, no. 2, pp. 223–234, 1975.

[77] Y. Yoshimoto, F. Matsumura, and N. Kamiya, "Simultaneous oscillations of ca2+ efflux and tension generation in the permealized plasmodial strand of physarum," *Cell motility*, vol. 1, no. 4, pp. 433–443, 1981.

[78] D. R. Farr, H. Amster, and M. Horisberger, "Composition and partial structure of the extracellular polysaccharide of physarum polycephalum," *Carbohydrate research*, vol. 24, no. 1, pp. 207–209, 1972.

[79] H. Sesaki and S. Ogihara, "Secretion of slime, the extracellular matrix of the plasmodium, as visualized with a fluorescent probe and its correlation with locomotion on the substratum," *Cell structure and function*, vol. 22, no. 2, pp. 279–289, 1997.

[80] C. R. Reid, M. Beekman, T. Latty, and A. Dussutour, "Amoeboid organism uses extracellular secretions to make smart foraging decisions," *Behavioral Ecology*, vol. 24, no. 4, pp. 812–818, 2013.

[81] C. R. Reid, T. Latty, A. Dussutour, and M. Beekman, "Slime mold uses an externalized spatial "memory" to navigate in complex environments," *Proceedings of the National Academy of Sciences*, vol. 109, no. 43, pp. 17 490–17 494, 2012.

[82] T. Ueda, T. Hirose, and Y. Kobatake, "Membrane biophysicsdf chemoreception and taxis in the plasmodium of physarum polycephalum," *Biophysical chemistry*, vol. 11, no. 3-4, pp. 461–473, 1980.

[83] C. R. Reid and T. Latty, "Collective behaviour and swarm intelligence in slime moulds," *FEMS microbiology reviews*, vol. 40, no. 6, K. Gibbs, Ed., pp. 798–806, Nov. 2016.

[84] A. Adamatzky, "Slime mold solves maze in one pass, assisted by gradient of chemoattractants," *IEEE Transactions on NanoBioscience*, vol. 11, no. 2, pp. 131–134, Jun. 2012.

[85] R. P. Boisseau, D. Vogel, and A. Dussutour, "Habituation in non-neural organisms: Evidence from slime moulds," *Proceedings of the Royal Society B: Biological Sciences*, vol. 283, no. 1829, p. 20 160 446, 2016.

[86] A. Dussutour, Q. Ma, and D. Sumpter, "Phenotypic variability predicts decision accuracy in unicellular organisms," *Proceedings of the Royal Society B*, vol. 286, no. 1896, p. 20 182 825, 2019.

[87] T. Latty and M. Beekman, "Slime moulds use heuristics based on within-patch experience to decide when to leave.," *The Journal of experimental biology*, vol. 218, no. Pt 8, pp. 1175–9, 2015.

[88] T. Nakagaki, H. Yamada, and T. Ueda, "Interaction between cell shape and contraction pattern in the physarum plasmodium," *Biophysical chemistry*, vol. 84, no. 3, pp. 195–204, 2000.

[89] C. R. Reid, S. Garnier, M. Beekman, and T. Latty, "Information integration and multiattribute decision making in non-neuronal organisms," *Animal Behaviour*, vol. 100, pp. 44–50, Feb. 2015.

[90] T. Saigusa, A. Tero, T. Nakagaki, and Y. Kuramoto, "Amoebae anticipate periodic events," *Physical review letters*, vol. 100, no. 1, p. 18 101, 2008.

[91] K. Terayama, T. Ueda, K. Kurihara, and Y. Kobatake, "Effect of sugars on salt reception in true slime moldphysarum polycephalum," *The Journal of membrane biology*, vol. 34, no. 1, pp. 369–381, 1977.

[92] M. J. Carlile, "Nutrition and chemotaxis in the myxomycete physarum polycephalum: The effect of carbohydrates on the plasmodium," *Microbiology*, vol. 63, no. 2, pp. 221–226, 1970.

[93] S. L. Stephenson and M. Schnittler, "Myxomycetes," *Handbook of the Protists*, pp. 1405–1431, 2017.

[94] D. Vogel and A. Dussutour, "Direct transfer of learned behaviour via cell fusion in non-neural organisms," *Proceedings of the Royal Society B: Biological Sciences*, vol. 283, no. 1845, p. 20 162 382, 2016.

[95] A. Dussutour, T. Latty, M. Beekman, and S. J. Simpson, "Amoeboid organism solves complex nutritional challenges," *Proceedings of the National Academy of Sciences*, vol. 107, no. 10, pp. 4607–4611, 2010.

[96] N. Dhanachandra, K. Manglem, and Y. J. Chanu, "Image segmentation using k -means clustering algorithm and subtractive clustering algorithm," *Procedia Computer Science*, vol. 54, pp. 764–771, Jan. 1, 2015.

[97] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu, "An efficient k-means clustering algorithm: Analysis and implementation,"

IEEE Transactions on Pattern Analysis & Machine Intelligence, no. 7, pp. 881–892, 2002.

[98] T. Therneau, "A package for survival analysis in s. version 2.38," 2015.

[99] D. Bates, D. Sarkar, M. D. Bates, and L. Matrix, "The lme4 package," *R package version*, vol. 2, no. 1, p. 74, 2007.

[100] A. Kuznetsova, P. B. Brockhoff, and R. H. B. Christensen, "Lmertest package: Tests in linear mixed effects models," *Journal of Statistical Software*, vol. 82, no. 13, 2017.

[101] K. Barton and M. K. Barton, "Package 'mumin'version 1," 2015.

[102] K.-i. Ueda, S. Takagi, Y. Nishiura, and T. Nakagaki, "Mathematical model for contemplative amoeboid locomotion," *Physical Review E*, vol. 83, no. 2, p. 21 916, 2011.

[103] J. Jones, "Characteristics of pattern formation and evolution in approximations of physarum transport networks," *Artificial Life*, vol. 16, no. 2, pp. 127–153, Apr. 2010.

[104] R. Halvorsrud and G. Wagner, "Growth patterns of the slime mold physarum on a nonuniform substrate," *Physical Review E*, vol. 57, no. 1, p. 941, 1998.

[105] T. Latty and M. Beekman, "Food quality affects search strategy in the acellular slime mould, physarum polycephalum," *Behavioral Ecology*, vol. 20, no. 6, pp. 1160–1167, 2009.

[106] J. W. Daniel and H. H. Baldwin, "Methods of culture for plasmodial myxomycetes," in *Methods in Cell Biology*, vol. 1, Elsevier, 1964, pp. 9–41.

[107] D. J. Knowles and M. J. Carlile, "The chemotactic response of plasmodia of the myxomycete physarum polycephalum to sugars and related compounds," *Microbiology*, vol. 108, no. 1, pp. 17–25, 1978.

[108] A. Takamatsu, E. Takaba, and G. Takizawa, "Environment-dependent morphology in plasmodium of true slime mold physarum polycephalum and a network growth model," *Journal of theoretical biology*, vol. 256, no. 1, pp. 29–44, 2009.

[109] A. Adamatzky, *Physarum Machines*, ser. World Scientific Series on Nonlinear Science Series A. WORLD SCIENTIFIC, Aug. 2010, vol. 74, ISBN: 978-981-4327-58-9.

[110] U. Nwodo, E. Green, and A. Okoh, "Bacterial exopolysaccharides: Functionality and prospects," *International journal of molecular sciences*, vol. 13, no. 11, pp. 14 002–14 015, 2012.

[111] J. J. McCormick, J. C. Blomquist, and H. P. Rusch, "Isolation and characterization of a galactosamine wall from spores and spherules of physarum polycephalum," *Journal of bacteriology*, vol. 104, no. 3, pp. 1119–1125, 1970.

[112] H. L. Simon and H. R. Henney, "Chemical composition of slime from three species of myxomycetes," *FEBS letters*, vol. 7, no. 1, pp. 80–82, 1970.

[113] M. Asgari and J. H. R. Henney, "Inhibition of growth and cell wall morphogenesis of bacillus subtilis by extracellular slime produced by physarum flavicomum.," *Cytobios*, vol. 20, no. 79-80, pp. 163–177, 1977.

[114] E. F. Haskins and A. A. Hinchee, "Light-and ultra-microscopical observations on the surface structure of the protoplasmodium, aphanoplasmodium, and phaneroplasmodium (myxomycètes)," *Canadian Journal of Botany*, vol. 52, no. 8, pp. 1835–1839, 1974.

[115] C. R. Reid and M. Beekman, "Solving the towers of hanoi - how an amoeboid organism efficiently constructs transport networks.," *The Journal of experimental biology*, vol. 216, pp. 1546–1551, 2013.

[116] A. Adamatzky, "Routing physarum with repellents," *The European Physical Journal E*, vol. 31, no. 4, pp. 403–410, Apr. 17, 2010.

[117] A. Adamatzky, "Physarum machines: Encapsulating reaction–diffusion to compute spanning tree," *Naturwissenschaften*, vol. 94, no. 12, pp. 975–980, Nov. 14, 2007.

[118] A. Adamatzky and J. Jones, "Towards physarum robots: Computing and manipulating on water surface," *Journal of Bionic Engineering*, vol. 5, no. 4, pp. 348–357, Dec. 1, 2008.

[119] A. Adamatzky, G. J. Martínez, V. S. Chapa-Vergara, R. Asomoza-Palacio, and C. R. Stephens, "Approximating mexican highways with slime mould," *Natural Computing*, vol. 10, no. 3, pp. 1195–1214, Sep. 28, 2011.

[120] A. SCHUMANN and A. ADAMATZKY, "Physarum spatial logic," *New Mathematics and Natural Computation*, vol. 07, no. 03, pp. 483–498, Sep. 21, 2011.

[121] T. Umedachi, K. Takeda, T. Nakagaki, R. Kobayashi, and A. Ishiguro, "Fully decentralized control of a soft-bodied robot inspired by true slime mold," *Biological Cybernetics*, vol. 102, no. 3, pp. 261–269, Mar. 4, 2010.

[122] T. W. Secomb and A. R. Pries, "The blood vasculature as an adaptive system: Role of mechanical sensing," in *Sensors and Sensing in Biology and Engineering*, Springer, 2003, pp. 187–196.

[123] A. R. Pries, B. Reglin, and T. W. Secomb, "Remodeling of blood vessels: Responses of diameter and wall thickness to hemodynamic and metabolic stimuli," *Hypertension*, vol. 46, no. 4, pp. 725–731, 2005.

[124] J. Vallverdú, O. Castro, R. Mayne, M. Talanov, M. Levin, F. Baluška, Y. Gunji, A. Dussutour, H. Zenil, and A. Adamatzky, "Slime mould: The fundamental mechanisms of biological cognition," *Biosystems*, vol. 165, pp. 57–70, 2018.

[125] A. Adamatzky, "Thirty seven things to do with live slime mould," in *Advances in Unconventional Computing*, Springer, 2017, pp. 709–738.

[126] C. Gao, C. Liu, D. Schenz, X. Li, Z. Zhang, M. Jusup, Z. Wang, M. Beekman, and T. Nakagaki, "Does being multi-headed make you better at solving problems? a survey of physarum-based models and computations," *Physics of life reviews*, vol. 29, pp. 1–26, 2019.

[127] T. Nakagaki and R. D. Guy, "Intelligent behaviors of amoeboid movement based on complex dynamics of soft matter," *Soft Matter*, vol. 4, no. 1, pp. 57–67, 2008.

[128] K. Alim, G. Amselem, F. Peaudecerf, M. P. Brenner, and A. Pringle, "Random network peristalsis in physarum polycephalum organizes fluid flows across an individual," *Proceedings of the National Academy of Sciences*, vol. 110, no. 33, pp. 13 306–13 311, 2013.

[129] O. L. Lewis, S. Zhang, R. D. Guy, and J. C. Del Alamo, "Coordination of contractility, adhesion and flow in migrating physarum amoebae," *Journal of The Royal Society Interface*, vol. 12, no. 106, p. 20 141 359, 2015.

[130] Y. Miyake, H. Tada, M. Yano, and H. Shimizu, "Relationship between intracellular period modulation and external environment change in physarum plasmodium," *Cell structure and function*, vol. 19, no. 6, pp. 363–370, 1994.

[131] K. Matsumoto, T. Ueda, and Y. Kobatake, "Propagation of phase wave in relation to tactic responses by the plasmodium of physarum polycephalum," *Journal of Theoretical Biology*, vol. 122, no. 3, pp. 339–345, 1986.

[132] R. L. Kincaid and T. E. Mansour, "Chemotaxis toward carbohydrates and amino acids in physarum polycephalum," *Experimental cell research*, vol. 116, no. 2, pp. 377–385, 1978.

[133] Y. Kagawa and A. Takamatsu, "Synchronization and spatiotemporal patterns in coupled phase oscillators on a weighted planar network," *Physical Review E*, vol. 79, no. 4, p. 046 216, 2009.

[134] M. Dirnberger, K. Mehlhorn, and T. Mehlhorn, "Introducing the slime mold graph repository," *Journal of Physics D: Applied Physics*, vol. 50, no. 26, 2017.

[135] M. Dirnberger, T. Kehl, and A. Neumann, "NEFI: Network Extraction from Images," *Scientific Reports*, vol. 5, pp. 1–10, 2015.

[136] M. D. Fricker, D. Akita, L. L. Heaton, N. Jones, B. Obara, and T. Nakagaki, "Automated analysis of Physarum network structure and dynamics," *Journal of Physics D: Applied Physics*, vol. 50, no. 25, 2017.

[137] W. Baumgarten, T. Ueda, and M. J. B. Hauser, "Plasmodial vein networks of the slime mold Physarum polycephalum form regular graphs," *Physical Review E*, vol. 82, no. 4, p. 046 113, Oct. 2010.

[138] M. Ito, R. Okamoto, and A. Takamatsu, "Characterization of adaptation by morphology in a planar biological network of plasmodial slime mold," *Journal of the Physical Society of Japan*, vol. 80, no. 7, pp. 1–7, 2011.

[139] M. Dirnberger and K. Mehlhorn, "Characterizing networks formed by P. polycephalum," *Journal of Physics D: Applied Physics*, vol. 50, no. 22, p. 224 002, Jun. 2017.

[140] W. Baumgarten and M. J. B. Hauser, "Functional organization of the vascular network of Physarum polycephalum," *Physical Biology*, vol. 10, no. 2, p. 026 003, Feb. 2013.

[141] K. Alim, N. Andrew, A. Pringle, and M. P. Brenner, "Mechanism of signal propagation in physarum polycephalum," *Proceedings of the National Academy of Sciences*, vol. 114, no. 20, pp. 5136–5141, 2017.

[142] A. Wacker and H. Gerhardt, "Endothelial development taking shape," *Current opinion in cell biology*, vol. 23, no. 6, pp. 676–685, 2011.

[143] D. J. Higham and N. J. Higham, *MATLAB guide*. SIAM, 2016.

[144] D. Bradley and G. Roth, "Adaptive Thresholding using the Integral Image," *Journal of Graphics Tools*, vol. 12, no. 2, pp. 13–21, Jan. 2007.

[145] F. Meyer, "Topographic distance and watershed lines," *Signal Processing*, vol. 38, no. 1, pp. 113–125, Jul. 1994.

[146] T. C. Lee, R. L. Kashyap, and C. N. Chu, "Building Skeleton Models via 3-D Medial Surface Axis Thinning Algorithms," *CVGIP: Graphical Models and Image Processing*, vol. 56, no. 6, pp. 462–478, Nov. 1994.

[147] C. R. Maurer, R. Qi, and V. Raghavan, "A linear time algorithm for computing exact Euclidean distance transforms of binary images in arbitrary dimensions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 2, pp. 265–270, Feb. 2003.

[148] M. Barthélemy, "Betweenness centrality in large complex networks," *The European Physical Journal B - Condensed Matter*, vol. 38, no. 2, pp. 163–168, Mar. 2004.

[149] L. Papadopoulos, P. Blinder, H. Ronellenfitsch, F. Klimm, E. Katifori, D. Kleinfeld, and D. S. Bassett, "Comparing two classes of biological distribution systems using network analysis," *PLoS Computational Biology*, vol. 14, pp. 1–31, 2018.

[150] J. Buhl, J. Gautrais, R. V. Solé, P. Kuntz, S. Valverde, J. L. Deneubourg, and G. Theraulaz, "Efficiency and robustness in ant networks of galleries," *European Physical Journal B*, vol. 42, no. 1, pp. 123–129, 2004.

[151] D. Bates, M. Maechler, B. Bolker, S. Walker, R. H. B. Christensen, H. Singmann, B. Dai, G. Grothendieck, P. Green, and M. B. Bolker, "Package 'lme4'," *Convergence*, vol. 12, no. 1, p. 2, 2015.

[152] R. Peterson, *Bestnormalize: Normalizing transformation functions, r package version 1.2. 0*, 2018.

[153] T. Nakagaki, M. Iima, T. Ueda, Y. Nishiura, T. Saigusa, A. Tero, R. Kobayashi, and K. Showalter, "Minimum-risk path finding by an adaptive amoebal network," *Physical review letters*, vol. 99, no. 6, p. 068 104, 2007.

[154] E. Katifori and M. O. Magnasco, "Quantifying loopy network architectures," *PloS one*, vol. 7, no. 6, e37994, 2012.

[155] E. Katifori, G. J. Szöllősi, and M. O. Magnasco, "Damage and fluctuations induce loops in optimal transport networks," *Physical review letters*, vol. 104, no. 4, p. 048 704, 2010.

[156] L. Papadopoulos, P. Blinder, H. Ronellenfitsch, F. Klimm, E. Katifori, D. Kleinfeld, and D. S. Bassett, "Comparing two classes of biological distribution systems using network analysis," *PLoS computational biology*, vol. 14, no. 9, e1006428, 2018.

[157] J. R. Banavar, A. Maritan, and A. Rinaldo, "Size and form in efficient transportation networks," *Nature*, vol. 399, no. 6732, pp. 130–132, May 1999.

[158] S. Navlakha, X. He, C. Faloutsos, and Z. Bar-Joseph, "Topological properties of robust biological and computational networks," *Journal of The Royal Society Interface*, vol. 11, no. 96, pp. 20 140 283–20 140 283, 2014.

[159] J. Jiang, X. Wang, and Y. C. Lai, "Optimizing biologically inspired transport networks by control," *Physical Review E*, vol. 100, no. 3, p. 32 309, 2019.

[160] T. L. Moore, J. S. Gulliver, L. Stack, and M. H. Simpson, "Stormwater management and climate change: Vulnerability and capacity for adaptation in urban and suburban contexts," *Climatic Change*, vol. 138, no. 3-4, pp. 491–504, 2016.

[161] P. F. Boulos, "Smart water network modeling for sustainable and resilient infrastructure," *Water Resources Management*, vol. 31, no. 10, pp. 3177–3188, 2017.

[162] M. Panteli, C. Pickering, S. Wilkinson, R. Dawson, and P. Mancarella, "Power system resilience to extreme weather: Fragility modeling, probabilistic impact assessment, and adaptation measures," *IEEE Transactions on Power Systems*, vol. 32, no. 5, pp. 3747–3757, 2017.

[163] C. Colman Meixner, F. Dikbiyik, M. Tornatore, C. Chuah, and B. Mukherjee, "Disaster-resilient virtual-network mapping and adaptation in optical networks," in *2013 17th International Conference on Optical Networking Design and Modeling (ONDM)*, 2013, pp. 107–112.

[164] X. He and E. J. Cha, "Modeling the damage and recovery of interdependent civil infrastructure network using dynamic integrated network model," *Sustainable and Resilient Infrastructure*, vol. 5, no. 3, pp. 152–167, 2020.

[165] J. Chen, C. Touati, and Q. Zhu, "A dynamic game approach to strategic design of secure and resilient infrastructure network," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 462–474, 2020.

[166] F. Flammini, A. Gaglione, N. Mazzocca, V. Moscato, and C. Pragliola, "Wireless sensor data fusion for critical infrastructure security," in *Proceedings of the International Workshop on Computational Intelligence in Security for Information Systems CISIS'08*, Springer, 2009, pp. 92–99.

[167] O. Linda, M. Manic, and T. R. McJunkin, "Anomaly detection for resilient control systems using fuzzy-neural data fusion engine," in *2011 4th International Symposium on Resilient Control Systems*, 2011, pp. 35–41.

[168] F. Rosas, K.-C. Chen, and D. Gündüz, "Social learning for resilient data fusion against data falsification attacks," *Computational social networks*, vol. 5, no. 1, p. 10, 2018.

[169] J. Briere, "Rapid restoration of critical infrastructures: An all-hazards paradigm for fusion centres," *International journal of critical infrastructures*, vol. 7, no. 1, pp. 21–36, 2011.

[170] A. Bejan and M. R. Errera, "Deterministic tree networks for fluid flow: Geometry for minimal flow resistance between a volume and one point," *Fractals*, vol. 05, no. 04, pp. 685–695, 1997.

[171] a. H. Reis, a. F. Miguel, and M. Aydin, "Constructal theory of flow architecture of the lungs.," *Medical physics*, vol. 31, no. 5, pp. 1135–1140, 2004.

[172] A. Bejan and S. Lorente, "Design with constructal theory," *Design with Constructal Theory*, no. September, pp. 1–529, 2008.

[173] a. H. Reis, "Constructal theory: From engineering to physics, and how flow systems develop shape and structure," *Applied Mechanics Reviews*, vol. 59, no. 5, p. 269, 2006.

[174] P. Winter and M. Zachariasen, "Euclidean steiner minimum trees: An improved exact algorithm," *Networks*, vol. 30, no. 3, pp. 149–166, 1997.

[175] T. Nakagaki, H. Yamada, and M. Hara, "Smart network solutions in an amoeboid organism," *Biophysical Chemistry*, vol. 107, no. 1, pp. 1–5, 2004.

[176] A. Tero, T. Nakagaki, K. Toyabe, K. Yumiki, and R. Kobayashi, "A method inspired by physarum for solving the steiner problem," *IJUC*, vol. 6, pp. 109–123, 2010.

[177] A. Tero, R. Kobayashi, and T. Nakagaki, "A mathematical model for adaptive transport network in path finding by true slime mold," *Journal of Theoretical Biology*, vol. 244, no. 4, pp. 553–564, 2007.

[178] H. Gao, T. Li, and L. Yang, "Quantitative determination of pore and throat parameters in tight oil reservoir using constant rate mercury intrusion technique," *Journal of Petroleum Exploration and Production Technology*, vol. 6, no. 2, pp. 309–318, 2016.

[179] P. Kenrick and C. Strullu-Derrien, "The origin and early evolution of roots," *PLANT PHYSIOLOGY*, vol. 166, no. 2, pp. 570–580, 2014.

[180] C. Kevin Boyce, "The evolutionary history of roots and leaves," in *Vascular Transport in Plants*, Elsevier, 2005, pp. 479–499, ISBN: 9780120884575.

[181] L. Sack and C. Scoffoni, "Leaf venation: Structure, function, development, evolution, ecology and applications in the past, present and future," *New Phytologist*, vol. 198, no. 4, pp. 983–1000, 2013.

[182] H. L. Sanders and S. E. Wyatt, "Leaf evolution and development: Advancing technologies, advancing understanding," *BioScience*, vol. 59, no. 1, pp. 17–26, 2009.

[183] H. Whitney, "A theorem on graphs," *The Annals of Mathematics*, vol. 32, no. 2, p. 378, Apr. 1931.

[184] R. Albert, H. Jeong, and A. L. Barabási, "Error and attack tolerance of complex networks," *Nature*, vol. 406, no. 6794, pp. 378–382, Jul. 27, 2000.

[185] M. Aldana and P. Cluzel, "A natural class of robust networks," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 15, pp. 8710–8714, Jul. 22, 2003.

[186] N. Barkai and S. Leibler, "Robustness in simple biochemical networks," *Nature*, vol. 387, no. 6636, pp. 913–917, Jun. 26, 1997.

[187] J. A. Draghi, T. L. Parsons, G. P. Wagner, and J. B. Plotkin, "Mutational robustness can facilitate adaptation," *Nature*, vol. 463, no. 7279, pp. 353–355, Jan. 21, 2010.

[188] M. E. Gáspár and P. Csermely, "Rigidity and flexibility of biological networks," *Briefings in Functional Genomics*, vol. 11, no. 6, pp. 443–456, Nov. 2012.

[189] R. S. Udan, J. C. Culver, and M. E. Dickinson, "Understanding vascular development," *Wiley Interdisciplinary Reviews: Developmental Biology*, vol. 2, no. 3, pp. 327–346, 2013.

[190] A. Chandrasekhar and S. Navlakha, "Neural arbors are pareto optimal," *Proceedings of the Royal Society B: Biological Sciences*, vol. 286, no. 1902, p. 20 182 727, 2019.

[191] A. Conn, V. U. Pedmale, J. Chory, and S. Navlakha, "High-resolution laser scanning reveals plant architectures that reflect universal network design principles," *Cell Systems*, vol. 5, no. 1, 53–62.e3, 2017.

[192] A. Hodge, G. Berta, C. Doussan, F. Merchan, and M. Crespi, "Plant root growth, architecture and function," *Plant and Soil*, vol. 321, no. 1-2, pp. 153–187, 2009.

[193] S. Katuwal, J. Vermang, W. M. Cornelis, D. Gabriels, P. Moldrup, and L. W. de Jonge, "Effect of root density on erosion and erodibility of a loamy soil under simulated rain," *Soil Science*, vol. 178, no. 1, pp. 29–36, 2013.

[194] W. R. Whalley, P. B. Leeds-Harrison, P. K. Leech, B. Riseley, and N. R. A. Bird, "The hydraulic properties of soil at root-soil interface," *Soil Science*, vol. 169, no. 2, pp. 90–99, Feb. 2004.

[195] H.-W. Yu, C.-H. Chen, and W.-C. Lo, "Mechanical behavior of the soil-root system," *Soil Science*, vol. 176, no. 2, pp. 99–109, 2011.

[196] D. P. Bebber, J. Hynes, P. R. Darrah, L. Boddy, and M. D. Fricker, "Biological solutions to transport network design," *Proceedings of the Royal Society B: Biological Sciences*, vol. 274, no. 1623, pp. 2307–2315, Sep. 2007.

[197] A. Bottinelli, v. E. Wilgenburg, D. J. T. Sumpter, and T. Latty, "Local cost minimization in ant transport networks: From small-scale data to large-scale trade-offs," *Journal of The Royal Society Interface*, vol. 12, no. 112, p. 20 150 780, 2015.

[198] G. Cabanes, E. Van Wilgenburg, M. Beekman, and T. Latty, "Ants build transportation networks that optimize cost and efficiency at the expense of robustness," *Behavioral Ecology*, vol. 26, no. 1, pp. 223–231, 2015.

[199] A. Perna, B. Granovskiy, S. Garnier, S. C. Nicolis, M. Labédan, G. Theraulaz, V. Fourcassié, and D. J. T. Sumpter, "Individual rules for trail pattern formation in argentine ants (linepithema humile)," *PLoS Computational Biology*, vol. 8, no. 7, B. Ermentrout, Ed., e1002592, 2012.

[200] E. van Wilgenburg and M. A. Elgar, "Colony structure and spatial distribution of food resources in the polydomous meat ant iridomyrmex purpureus," *Insectes Sociaux*, vol. 54, no. 1, pp. 5–10, 2007.

[201] J. Buhl, J. Gautrais, V. R. Sole, P. Kuntz, S. Valverde, J. L. Deneubourg, and G. Theraulaz, "Efficiency and robustness in ant networks of galleries," *The European Physical Journal B*, vol. 42, no. 1, pp. 123–129, 2004.

[202] K. Socha and M. Dorigo, "Ant colony optimization for continuous domains," *European Journal of Operational Research*, vol. 185, pp. 758–762, 2006.

[203] M. S. Davies and J. Blackwell, "Energy saving through trail following in a marine snail," *Proceedings of the Royal Society B: Biological Sciences*, vol. 274, no. 1614, pp. 1233–1236, 2007.

[204] A. Perna and T. Latty, "Animal transportation networks," *Journal of The Royal Society Interface*, vol. 11, no. 100, p. 20 140 334, 2014.

[205] K. Oka, S. Aoyagi, Y. Arai, Y. Isono, G. Hashiguchi, and H. Fujita, "Fabrication of a micro needle for a trace blood test," *Sensors and Actuators A: Physical*, vol. 97-98, pp. 478–485, 2002.

[206] L. Chen, A. L. DeVries, and C. H. Cheng, "Convergent evolution of antifreeze glycoproteins in Antarctic notothenioid fish and Arctic cod.," *Proceedings of the*

National Academy of Sciences of the United States of America, vol. 94, no. 8, pp. 3817–22, 1997.

[207] F. E. Fish, P. W. Weber, M. M. Murray, and L. E. Howle, "The tubercles on humpback whales' flippers: Application of bio-inspired technology," *Integrative and Comparative Biology*, vol. 51, no. 1, pp. 203–213, 2011.

[208] L. Ninan, J. Monahan, R. L. Stroshine, J. J. Wilker, and R. Shi, "Adhesive strength of marine mussel extracts on porcine skin," *Biomaterials*, vol. 24, no. 22, pp. 4091–4099, 2003.

[209] Z. W. Geem, "Particle-swarm harmony search for water network design," *Engineering Optimization*, vol. 41, no. 4, pp. 297–311, 2009.

[210] H. Zhou, Z. Zhang, Y. Wu, and T. Qian, "Bio-inspired dynamic composition and reconfiguration of service-oriented internetware systems," in, Springer, Berlin, Heidelberg, 2011, pp. 364–373.

[211] Z. Yang, B. Yu, and C. Cheng, "A parallel ant colony algorithm for bus network optimization," *Computer-Aided Civil and Infrastructure Engineering*, vol. 22, no. 1, pp. 44–55, 2007.

[212] T. Nelson and N. Dengler, "Leaf vascular pattern formation.," *The Plant Cell*, vol. 9, no. 7, pp. 1121–1135, 1997.

[213] H. Ronellenfitsch and E. Katifori, "Global optimization, local adaptation, and the role of growth in distribution networks," *Physical Review Letters*, vol. 117, no. 13, pp. 1–5, 2016.

[214] A. Roth-Nebelsick, "Evolution and function of leaf venation architecture: A review," *Annals of Botany*, vol. 87, no. 5, pp. 553–566, May 2001.

[215] W. S. Judd, C. S. Campbell, E. A. Kellogg, P. F. Stevens, and M. J. Donoghue, *Plant systematics*. Sinauer Sunderland, Massachusetts, USA, 2002.

[216] A. Runions, M. Fuhrer, B. Lane, P. Federl, A.-G. Rolland-Lagan, and P. Prusinkiewicz, "Modeling and visualization of leaf venation patterns," *ACM Transactions on Graphics*, vol. 24, no. 3, p. 702, 2005.

[217] T. Sachs, "The control of the patterned differentiation of vascular tissues," *Advances in Botanical Research*, vol. 9, pp. 151–262, 1981.

[218] R. Aloni, K. Schwalm, M. Langhans, and C. I. Ullrich, "Gradual shifts in sites of free-auxin production during leaf-primordium development and their role in vascu-

lar differentiation and leaf morphogenesis in arabidopsis," *Planta*, vol. 216, no. 5, pp. 841–853, 0.

[219] C. D. Murray, "The physiological principle of minimum work: I. the vascular system and the cost of blood volume.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 12, no. 3, pp. 207–14, Mar. 1926.

[220] J. Teich, "Pareto-front exploration with uncertain objectives," in, Springer, Berlin, Heidelberg, 2001, pp. 314–328.

[221] G. Robins and A. Zelikovsky, "Minimum steiner tree construction," *Handbook of Algorithms for VLSI Physical Automation*, pp. 487–508, 2008.

[222] J. A. Wald and C. J. Colbourn, "Steiner trees in probabilistic networks," *Microelectronics Reliability*, vol. 23, no. 5, pp. 837–840, 1983.

[223] G. Ajwani, C. Chu, and W. K. Mak, "Foars: Flute based obstacle-avoiding rectilinear steiner tree construction," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 30, no. 2, pp. 194–204, 2011.

[224] W. K. Chow, L. Li, E. F. Y. Young, and C. W. Sham, "Obstacle-avoiding rectilinear steiner tree construction in sequential and parallel approach," *Integration, the VLSI Journal*, vol. 47, no. 1, pp. 105–114, 2014.

[225] I. Ljubic, R. Weiskircher, U. Pferschy, G. W. Klau, P. Mutzel, and M. Fischetti, "An algorithm framework for the exact solution of the prize-collecting steiner tree problem," pp. 1–25, 2005.

[226] M. Müller-Hannemann and S. Tazari, "A near linear time approximation scheme for steiner tree among obstacles in the plane," *Computational Geometry: Theory and Applications*, vol. 43, no. 4, pp. 395–409, 2010.

[227] M. Fampa, J. Lee, and N. Maculan, "An overview of exact algorithms for the euclidean steiner tree problem in n -space," *International Transactions in Operational Research*, vol. 23, no. 5, pp. 861–874, 2016.

[228] C.-H. L. C.-H. Liuf, S.-Y. Y. S.-Y. Yuan, S.-Y. K. S.-Y. Kuo, and Y.-H. C. Y.-H. Chou, "An o(n log n) path-based obstacle-avoiding algorithm for rectilinear steiner tree construction," *2009 46th ACM/IEEE Design Automation Conference*, pp. 314–319, 2009.

[229] R. Fonseca, M. Brazil, P. Winter, and M. Zachariasen, "Faster exact algorithms for computing steiner trees in higher dimensional euclidean spaces," *11th DIMACS Implementation Challenge. Brown University*, 2014.

[230] J. Y. Suen and S. Navlakha, "Travel in city road networks follows similar transport trade-off principles to neural and plant arbors," *Journal of the Royal Society Interface*, vol. 16, no. 154, May 1, 2019.

[231] M. Willebeek-LeMair and A. Reeves, "Strategies for dynamic load balancing on highly parallel computers," *IEEE Transactions on Parallel and Distributed Systems*, vol. 4, no. 9, pp. 979–993, 1993.

[232] M. Barthelemy, "Betweenness centrality in large complex networks," *The European Physical Journal B - Condensed Matter*, vol. 38, no. 2, pp. 163–168, Mar. 1, 2004.

[233] M. J. Newman, "A measure of betweenness centrality based on random walks," *Social Networks*, vol. 27, no. 1, pp. 39–54, 2005.

[234] United States Census Bureau, *2010 Decennial Census*, 2010.

[235] S. Biba, K. M. Curtin, and G. Manca, "A new method for determining the population with walking access to transit," *International Journal of Geographical Information Science*, vol. 24, no. 3, pp. 347–364, 2010.

[236] D. O'Sullivan, A. Morrison, and J. Shearer, "Using desktop gis for the investigation of accessibility by public transport: An isochrone approach," *International Journal of Geographical Information Science*, vol. 14, no. 1, pp. 85–104, 2000.

[237] J. H. Atkinson and D. M. Potts, "Stability of a shallow circular tunnel in cohesionless soil," *Géotechnique*, vol. 27, no. 2, pp. 203–215, Jun. 1977.

[238] K. Wong, C. Ng, Y. Chen, and X. Bian, "Centrifuge and numerical investigation of passive failure of tunnel face in sand," *Tunnelling and Underground Space Technology*, vol. 28, no. 1, pp. 297–303, Mar. 2012.

[239] S. F. Timoshenko and J. N. Goodier, *Theory of Elasticity*, Second Edi. McGraw-Hill, 1951, p. 499.

[240] N. Muskhelishvili, *Some basic problems of the mathematical theory of elasticity*. Springer science + Business media dordrecht, 1977, p. 746, ISBN: 9789048182459.

[241] R. Hill, *The Mathematical theory of plasticity*. New York, New York, USA: Oxford University Press, 1950, p. 354.

[242] Z. Nehari, *Conformal Mapping*. McGraw-Hill, 1952.

[243] H.-S. Yu, *Cavity Expansion Methods in Geomechanics*. Dordrecht: Springer Netherlands, 2000, p. 385, ISBN: 978-90-481-4023-7.

[244] E. Andò, "Experimental investigation of microstructural changes in deforming granular media using x-ray tomography," PhD thesis, Université de Grenoble, 2013.

[245] J. Desrues and G. Viggiani, "Strain localization in sand: an overview of the experimental results obtained in Grenoble using stereophotogrammetry," *International Journal for Numerical and Analytical Methods in Geomechanics*, vol. 28, no. 4, pp. 279–321, Apr. 2004.

[246] O. Stamati, "Impact of meso-scale heterogeneities on the mechanical behaviour of concrete : insights from in-situ x-ray tomography and E-FEM modelling," Theses, Université Grenoble Alpes [2020-....], May 2020.

[247] O. Stamati, E. Andò, E. Roubin, R. Cailletaud, M. Wiebicke, G. Pinzon, C. Couture, R. C. Hurley, R. Caulk, D. Caillerie, *et al.*, "Spam: Software for practical analysis of materials," *Journal of Open Source Software*, vol. 5, no. 51, p. 2286, 2020.

[248] h. L.-h. Zhao, Y.-g. gao Tan, S.-h. hong Hu, D.-p. ping Deng, and X.-p. ping Yang, "Upper bound analysis of ultimate pullout capacity of shallow 3-D circular plate anchors based on nonlinear Mohr-Coulomb failure criterion," *Journal of Central South University*, vol. 25, no. 9, pp. 2272–2288, Sep. 2018.

[249] C. Y. Cheuk, D. J. White, and M. D. Bolton, "Uplift Mechanisms of Pipes Buried in Sand," *Journal of Geotechnical and Geoenvironmental Engineering*, vol. 134, no. 2, pp. 154–163, Feb. 2008.

[250] R. Zhang and X. Yang, "Limit analysis of active and passive mechanisms of shallow tunnels in nonassociative soil with changing water table," *International Journal of Geomechanics*, vol. 18, no. 7, pp. 1–13, 2018.

[251] H. B. Mühlhaus and I. Vardoulakis, "The thickness of shear bands in granular materials," *Géotechnique*, vol. 37, no. 3, pp. 271–283, Sep. 1987.

[252] P. A. Vermeer, "The orientation of shear bands in biaxial tests," *Geotechnique*, vol. 40, no. 2, pp. 223–236, Jun. 1990.

[253] D. J. White, A. J. Barefoot, and M. D. Bolton, "Centrifuge modelling of upheaval buckling in sand," *International Journal of Physical Modelling in Geotechnics*, vol. 1, no. 2, pp. 19–28, Jun. 2001.

[254] D. J. White, C. Y. Cheuk, and M. D. Bolton, "The uplift resistance of pipes and plate anchors buried in sand," *Geotechnique*, vol. 58, no. 10, pp. 771–779, 2008.

[255] M. Oda and H. Kazama, "Microstructure of shear bands and its relation to the mechanisms of dilatancy and failure of dense granular soils," *Géotechnique*, vol. 48, no. 4, pp. 465–481, Aug. 1998.

[256] K. A. Alshibli and A. Hasan, "Spatial variation of void ratio and shear band thickness in sand using X-ray computed tomography," *Géotechnique*, vol. 58, no. 4, pp. 249–257, Apr. 2008.

[257] K. H. Roscoe, "The Influence of Strains in Soil Mechanics," *Géotechnique*, vol. 20, no. 2, pp. 129–170, Jun. 1970.

[258] J. Desrues, R. Chambon, M. Mokni, and F. Mazerolle, "Void ratio evolution inside shear bands in triaxial sand specimens studied by computed tomography," *Géotechnique*, vol. 46, no. 3, pp. 529–546, Sep. 1996.

[259] T. M. Evans and J. D. Frost, "Multiscale investigation of shear bands in sand: Physical and numerical experiments," *International Journal for Numerical and Analytical Methods in Geomechanics*, vol. 34, no. 15, n/a–n/a, Oct. 2010.

[260] G. A. Narsilio and J. C. Santamarina, "Terminal densities," *Geotechnique*, vol. 58, no. 8, pp. 669–674, 2008.

[261] K. Miura, S. Miura, and S. Toki, "Deformation Behavior of Anisotropic Dense Sand Under Principal Stress Axes Rotation," *Soils and Foundations*, vol. 26, no. 1, pp. 36–52, Mar. 1986.

[262] Y. Yang and H. Yu, "Finite element analysis of anchor plates using non-coaxial models," *Journal of Rock Mechanics and Geotechnical Engineering*, vol. 2, no. 2, pp. 178–187, May 2010.

[263] Z. Chen and M. Huang, "Non-coaxial behavior modeling of sands subjected to principal stress rotation," *Acta Geotechnica*, vol. 15, no. 3, pp. 655–669, 2020.

[264] K. Ishihara and I. Towhata, "Sand Response to Cyclic Rotation of Principal Stress Directions as Induced by Wave Loads," *Soils and Foundations*, vol. 23, no. 4, pp. 11–26, Dec. 1983.

[265] Y. Cai, H.-S. Yu, D. Wanatowski, and X. Li, "Noncoaxial Behavior of Sand under Various Stress Paths," *Journal of Geotechnical and Geoenvironmental Engineering*, vol. 139, no. 8, pp. 1381–1395, Aug. 2013.

[266] R. L. Handy, "The Arch in Soil Arching," *Journal of Geotechnical Engineering*, vol. 111, no. 3, pp. 302–318, Mar. 1985.

[267] K. Harrop-Williams, "Arch in Soil Arching," *Journal of Geotechnical Engineering*, vol. 115, no. 3, pp. 415–419, Mar. 1989.

[268] B. Chevalier, G. Combe, and P. Villard, "Experimental and discrete element modeling studies of the trapdoor problem: Influence of the macro-mechanical frictional parameters," *Acta Geotechnica*, vol. 7, no. 1, pp. 15–39, 2012.

[269] P. Guo and S. Zhou, "Arch in granular materials as a free surface problem," *International Journal for Numerical and Analytical Methods in Geomechanics*, vol. 37, no. 9, pp. 1048–1065, Jun. 2013.

[270] R. S. Dalvi and P. J. Pise, "Analysis of Arching in Soil-Passive State," *Indian Geotechnical Journal*, vol. 42, no. 2, pp. 106–112, Jun. 2012.

[271] R. C. K. Wong and P. K. Kaiser, "Performance Assessment of Tunnels in Cohesionless Soils," *Journal of Geotechnical Engineering*, vol. 117, no. 12, pp. 1880–1901, Dec. 1991.

[272] F. Patino-Ramirez and C. Arson, "Numerical analysis of the plastic zone in drained soil around a pressurized cylindrical cavity under biaxial stress," in *IS Atlanta: International Symposium in Geo-Mechanics from Micro to Macro*, Atlanta, GA, 2018.

[273] F. X. Diebold, "A Personal Perspective on the Origin(s) and Development of 'Big Data': The Phenomenon, the Term, and the Discipline, Second Version," *SSRN Electronic Journal*, 2013.

[274] F. Provost and T. Fawcett, "Data Science and its Relationship to Big Data and Data-Driven Decision Making," *Big Data*, vol. 1, no. 1, pp. 51–59, Mar. 2013.

[275] W. van der Aalst, "Data Science in Action," in *Process Mining*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2016, pp. 3–23.

[276] M. A. Shahin, M. B. Jaksa, and H. R. Maier, "Recent Advances and Future Challenges for Artificial Neural Systems in Geotechnical Engineering Applications," *Advances in Artificial Neural Systems*, vol. 2009, pp. 1–9, 2009.

[277] N. Puri, H. D. Prasad, and A. Jain, "Prediction of Geotechnical Parameters Using Machine Learning Techniques," *Procedia Computer Science*, vol. 125, pp. 509–517, Jan. 2018.

[278] H. Zhou, H. Liu, G. Kong, and Z. Cao, "Analytical solution for pressure-controlled elliptical cavity expansion in elastic–perfectly plastic soil," *Géotechnique Letters*, vol. 4, no. 2, pp. 72–78, Apr. 2014.

[279] S. Wang, L. Sun, A. Au, T. Yang, and C. Tang, "2D-numerical analysis of hydraulic fracturing in heterogeneous geo-materials," *Construction and Building Materials*, vol. 23, no. 6, pp. 2196–2206, Jun. 2009.

[280] B. Shu, "Numerical Modeling Evaluation of Drilling Mud Loss Risk for a HDD Borehole in a Low-Strength Competent Bedrock," *Journal of Pipeline Systems Engineering and Practice*, vol. 9, no. 4, p. 04 018 011, Nov. 2018.

[281] M. E. Baumert, E. N. Allouche, and I. D. Moore, "Drilling fluid considerations in design of engineered horizontal directional drilling installations," *International Journal of Geomechanics*, vol. 5, no. 4, pp. 339–349, Dec. 2005.

[282] ASTM International, *F1962-11 Standard Guide for Use of Maxi-Horizontal Directional Drilling for Placement of Polyethylene Pipe or Conduit Under Obstacles, Including River Crossings*, West Conshohocken, PA, 2011.

[283] M. J. Kennedy, G. D. Skinner, and I. D. Moore, "Limiting Slurry Pressures to Control Hydraulic Fracturing in Directional Drilling Operations in Purely Cohesive Soil," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1976, no. 1, pp. 172–180, Jan. 2006.

[284] B. Shu, S. Zhang, and M. Liang, "Estimation of the maximum allowable drilling mud pressure for a horizontal directional drilling borehole in fractured rock mass," *Tunnelling and Underground Space Technology*, vol. 72, pp. 64–72, Feb. 2018.

[285] H. Xia, "Investigation of maximum mud pressure within sand and clay during horizontal directional drilling," PhD thesis, Queen's University, 2009, p. 272.

[286] F. H. Kulhawy and P. W. Mayne, "Manual on Estimating Soil Properties for Foundation Design," Tech. Rep., Aug. 1990, p. 299.

[287] B. Ladanyi and B. Hoyaux, "A study of the trap-door problem in a granular mass," *Canadian Geotechnical Journal*, vol. 6, no. 1, pp. 1–14, Feb. 1969.

[288] A. Rostami, Y. Yi, and A. Bayat, "Estimation of maximum annular pressure during hdd in noncohesive soils," *International Journal of Geomechanics*, vol. 17, no. 4, 2017.

[289] E. N. Allouche, S. T. Ariaratnam, and J. S. Lueke, "Horizontal directional drilling: Profile of an emerging industry," *Journal of Construction Engineering and Management*, vol. 126, no. 1, pp. 68–76, 2000.

[290] M. Sarireh, M. Najafi, and L. Slavin, "Usage and applications of horizontal directional drilling," Reston, VA: American Society of Civil Engineers, 2012, pp. 1835–1847, ISBN: 9780784412619.

[291] D. Bennett and S. T. Ariaratnam, *Horizontal Directional Drilling: Good Practices Guidelines*, 4th. North American Society for Trenchless Technology, 2017.

[292] X. Yan, S. T. Ariaratnam, S. Dong, and C. Zeng, "Horizontal directional drilling: State-of-the-art review of theory and applications," *Tunnelling and Underground Space Technology*, vol. 72, pp. 162–173, Feb. 1, 2018.

[293] M. Rabiei, Y. Yi, A. Bayat, and R. Cheng, "General method for pullback force estimation for polyethylene pipes in horizontal directional drilling," *Journal of Pipeline Systems Engineering and Practice*, vol. 7, no. 3, Aug. 1, 2016.

[294] B. Shu, S. Zhang, and M. Liang, "Estimation of the maximum allowable drilling mud pressure for a horizontal directional drilling borehole in fractured rock mass," *Tunnelling and Underground Space Technology*, vol. 72, pp. 64–72, Feb. 1, 2018.

[295] G. Duyvestyn, "Comparison of predicted and observed hdd installation loads for various calculation methods," 2009.

[296] M. Kennedy, G. Skinner, and I. Moore, "Elastic calculations of limiting mud pressures to control hydrofracturing during hdd," 2004, pp. 1–11.

[297] D. Huey, J. Hair, and K. McLeod, "Installation loading and stress analysis involved with pipelines installed by horizontal directional drilling," North American Society for Trenchless Technology, Chicago, IL (United States), Aug. 1, 1996.

[298] A. S. for Testing and Materials, "F1962-11 standard guide for use of maxi-horizontal directional drilling for placement of polyethylene pipe or conduit under obstacles, including river crossings," vol. 04, pp. 1–18, 2011.

[299] D. A. Holway and T. J. Case, "Mechanisms of dispersed central-place foraging in polydomous colonies of the argentine ant," *Animal Behaviour*, vol. 59, no. 2, pp. 433–441, 2000.

[300] A. C. Kakas, D. Cohn, S. Dasgupta, A. G. Barto, G. A. Carpenter, S. Grossberg, G. I. Webb, M. Dorigo, M. Birattari, H. Toivonen, J. Timmis, J. Branke, H. Toivonen, A. L. Strehl, C. Drummond, A. Coates, P. Abbeel, A. Y. Ng, F. Zheng, G. I. Webb, and P. Tadepalli, "Ant colony optimization," in *Encyclopedia of Machine Learning*, Boston, MA: Springer US, 2011, pp. 36–39.

[301] M. Dorigo and G. Di Caro, "Ant colony optimization: A new meta-heuristic," vol. 2, IEEE, 1999, pp. 1470–1477, ISBN: 0-7803-5536-9.

[302] M. Dorigo and T. Stützle, "Ant colony optimization: Overview and recent advances," in *International Series in Operations Research and Management Science*, vol. 272, Springer New York LLC, 2019, pp. 311–351.

[303] M. Samà, P. Pellegrini, A. D'Ariano, J. Rodriguez, and D. Pacciarelli, "Ant colony optimization for the real-time train routing selection problem," *Transportation Research Part B: Methodological*, vol. 85, pp. 89–108, 2016.

[304] T. Liao, K. Socha, M. A. De Oca, T. Stutzle, and M. Dorigo, "Ant colony optimization for mixed-variable optimization problems," *IEEE Transactions on Evolutionary Computation*, vol. 18, no. 4, pp. 503–518, 2014.